

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 25, 2006, 07:55:45 ; Search time 188 Seconds
(without alignments)
535.201 Million cell updates/sec

Title: US-09-942-052A-728

Perfect score: 1198

Sequence: 1 MAAPQLRRSRCAVPPRGDF.....LKSLMKILSEVTPDQSKPEN 229

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_21:*

- 1: Geneseq1980s:*
- 2: Geneseq1990s:*
- 3: Geneseq2000s:*
- 4: Geneseq2001s:*
- 5: Geneseq2002s:*
- 6: Geneseq2003as:*
- 7: Geneseq2003bs:*
- 8: Geneseq2004s:*
- 9: Geneseq2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1198	100.0	229	4	AAM40223 Human pol
2	1198	100.0	229	5	ABG34856 Human can
3	1198	100.0	221	4	AAM42009 Human pol
4	138.5	11.6	164	5	ABG34855 Human cdn
5	113	9.4	233	4	AAM93668 Human pol
6	113	9.4	233	4	AAG93290 Human pol
7	113	9.4	233	7	Adj69163 Human hea
8	113	9.4	233	8	ADL31517 Human pro
9	92	7.7	695	5	ABBS4167 Human pro
10	92	7.7	695	8	ABBS4167 Human pro
11	90	7.5	269	4	AAO08972 Bacterial
12	89	7.4	572	8	ABM83354 Human dia
13	89	7.4	623	4	AAB93182 Human pro
14	89	7.4	623	5	ABBS97233 Novel hum
15	89	7.4	652	4	AAB93182 Human pro
16	89	7.4	664	4	AAB83843 Human pro
17	89	7.4	664	4	AAB83843 Amino aci
18	88.5	7.4	388	8	ADQ08702 Human Chf
19	87.5	7.3	1245	7	ADQ08702 Ciona int
20	86.5	7.2	470	8	Adc32818 Human nov
21	86.5	7.2	669	6	ABM83305 Human dia
22	86	7.2	158	7	ABU23878 Protein e
23	85.5	7.1	571	8	Adh88711 Enterococ
24	84.5	7.1	485	5	ABP51336 Human MDD

25	84	7.0	496	4	AAB95788 Human pro
26	84	7.0	496	7	ADC31196 Human nov
27	83.5	7.0	479	4	AAU39678 Proptonib
28	83.5	7.0	479	6	ABM36197 Proptonib
29	83	6.9	598	2	AAM59461 Microbial
30	83	6.9	598	2	AAM59463 Microbial
31	83	6.9	598	2	AAM59457 Microbial
32	83	6.9	598	2	AAM59458 Microbial
33	83	6.9	598	2	AAM59459 Microbial
34	83	6.9	598	2	AAM59462 Microbial
35	83	6.9	598	2	AAM59460 Microbial
36	83	6.9	598	2	AAM59464 Microbial
37	83	6.9	599	2	AAM59496 Microbial
38	83	6.9	606	1	AAP90110 Microbial
39	83	6.9	607	2	AAM59456 Microbial
40	83	6.9	607	2	AAM59465 Microbial
41	83	6.9	663	6	ABU23749 Protein e
42	82.5	6.9	278	8	ADX72944 Plant ful
43	82.5	6.9	325	3	AAV90242 Human cyc
44	82.5	6.9	493	8	ADN72405 Thale cre
45	82.5	6.9	493	9	AEA16977 Arabidops

ALIGNMENTS

RESULT 1
AAM40223

ID AAM40223 standard; protein; 229 AA.

AC AAM40223;

DT 22-OCT-2001 (first entry)

DE Human polypeptide SEQ ID NO 3368.

KW Human; nontropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia.

OS Homo sapiens.

XX WO200153312-A1.

XX 26-JUL-2001.

XX 26-DEC-2000; 2000WO-US034263.

XX 23-DEC-1999; 99US-00471275.

XX 21-JAN-2000; 2000US-00488725.

XX 25-APR-2000; 2000US-00523117.

XX 20-JUN-2000; 2000US-00598042.

XX 19-JUL-2000; 2000US-00620312.

XX 03-AUG-2000; 2000US-00653450.

XX 14-SEP-2000; 2000US-00662191.

XX 29-NOV-2000; 2000US-00693036.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;

XX Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;

XX Zhou P, Goodrich R, Drmanac RT;

XX WPI; 2001-442253/47.

XX N-PSDB; AAI59379.

XX Novel nucleic acids and polypeptides, useful for treating disorders such as central nervous system injuries.

Example 5; SEQ ID NO 3368; 10078pp; English.

PS The invention relates to human nucleic acids (AA157799-AA161369) and the
XX encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders. Note: The sequence data for this patent did not form
CC part of the printed specification
XX
SQ Sequence 229 AA;

Query Match 100.0%; Score 1198; DB 4; Length 229;
Best Local Similarity 100.0%; Pred. No. 9.7e-127;
Matches 229; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAAQPLRHRSRCATPPRGDFCGGTERAIDQASFTTSMEDTQVVGKSSPLGPAIGAEAP 60
DB 1 MAAQPLRHRSRCATPPRGDFCGGTERAIDQASFTTSMEDTQVVGKSSPLGPAIGAEAP 60

QY 61 AAGPOLPSWLQPERCAVFOCAQCHAVLADSLSRSLGAVVFSRVNNVLEAPFL 120
DB 61 AAGPOLPSWLQPERCAVFOCAQCHAVLADSLSRSLGAVVFSRVNNVLEAPFL 120

QY 121 VGIEGSLKSTYNLLFCGSGIPVGFHLYSTHAALALRGHFCCLSSDKMVCYLLKTKAIV 180
DB 121 VGIEGSLKSTYNLLFCGSGIPVGFHLYSTHAALALRGHFCCLSSDKMVCYLLKTKAIV 180

QY 181 NASEMDIQNVPLSEKIAELKEKIVLTNRKLSLMLKILSEVTPDOSKPEN 229
DB 181 NASEMDIQNVPLSEKIAELKEKIVLTNRKLSLMLKILSEVTPDOSKPEN 229

RESULT 2
ABG34856
ID ABG34856 standard; protein; 229 AA.
AC ABG34856;
XX
XX 15-JUL-2002 (first entry)
XX Human cancer related protein encoded by cDNA 85PIB3.
XX Human; cytostatic; 85PIB3; cancer; immunogen; chromosome 15q14.
XX Homo sapiens.
XX WO200218578-A2.
XX
XX 07-MAR-2002.
XX
XX 28-AUG-2001; 2001WO-US026838.
XX
XX 28-AUG-2000; 2000US-0228432P.
XX
XX (AGEN-) AGENSYS INC.
XX Raitano AB, Faris M, Hubert RS, Afar D, Ge W, Challita-Eid P;
PI Jakobovits A;
XX
XX WPI: 2002-382963/41.
DR N-PSDB; ABK70506.
XX
XX Composition for modulating the status of 85PIB3 protein or a molecule
PT comprising a substance e.g. antibody specific to, nucleic acid encoding,

or ribozyme of 85PIB3.

Claim 34; Fig 2; 201pp; English.

The invention relates to a composition comprising a substance that modulate the status of 85PIB3, where the status of a cell expresses 85PIB3 gene product is modulated. Also included are a composition comprising a peptide region of 5 amino acids of the 85PIB3 protein, in any whole number increment up to 229 that includes an aa position selected from an aa position having a value greater than 0.5 in the hydrophilicity profile, an aa position having a value less than 0.5 in the hydrophobicity profile, an aa position having a value greater than 0.5 in the percent accessible residue profile, an aa position having a value greater than 0.5 in the average flexibility profile, or an aa position having a value greater than 0.5 in the beta-turn profile; a polynucleotide that encodes analogue peptide of 8, 9, 10 or 11 contiguous residues of the 85PIB3 protein; a recombinant protein comprising the antigen-binding region of a monoclonal antibody; a non-human transgenic animal that produces an antibody that binds to the 85PIB3 protein; a hybridoma that produces antibody specific to the protein; a single chain monoclonal antibody (MAB) that comprises the variable domains of the heavy and monoclonal antibodies specific to the protein; a vector comprising a polynucleotide that encodes the MAB; inhibiting growth of cancer cells or treating a patient who bears cancer cells that expresses the protein, by administering the protein, antibody, polynucleotide encoding the protein, antisense polynucleotide to the polynucleotide, ribozyme that cleaves the polynucleotide and T cells that specifically recognize the protein; and generating a mammalian immune response directed to the protein exposing cells of the mammal's immune system to an immunogenic portion of the protein or polynucleotide. The composition, which comprises an antibody specific to the protein, is useful for delivering a cytotoxic agent to a cell that expresses the protein by providing a cytotoxic agent conjugated to antibody and exposing the cell to the antibody-agent conjugate. The methods are useful for inhibiting growth of cancer cells or treating a patient who bears cancer cells that expresses the protein, for generating a mammalian immune response directed to the protein, for detecting the presence of the protein or polynucleotide in a biological sample in a patient who has or who is suspected of having cancer and for monitoring 85PIB3 in a biological sample from a patient who has or who is suspected of having cancer. The gene for 85PIB3 is located on human chromosome 15q14. The present sequence is the 85PIB3 protein

Query Match 100.0%; Score 1198; DB 5; Length 229;
Best Local Similarity 100.0%; Pred. No. 9.7e-127;
Matches 229; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAAQPLRHRSRCATPPRGDFCGGTERAIDQASFTTSMEDTQVVGKSSPLGPAIGAEAP 60
DB 1 MAAQPLRHRSRCATPPRGDFCGGTERAIDQASFTTSMEDTQVVGKSSPLGPAIGAEAP 60

QY 61 AAGPOLPSWLQPERCAVFOCAQCHAVLADSLSRSLGAVVFSRVNNVLEAPFL 120
DB 61 AAGPOLPSWLQPERCAVFOCAQCHAVLADSLSRSLGAVVFSRVNNVLEAPFL 120

QY 121 VGIEGSLKSTYNLLFCGSGIPVGFHLYSTHAALALRGHFCCLSSDKMVCYLLKTKAIV 180
DB 121 VGIEGSLKSTYNLLFCGSGIPVGFHLYSTHAALALRGHFCCLSSDKMVCYLLKTKAIV 180

QY 181 NASEMDIQNVPLSEKIAELKEKIVLTNRKLSLMLKILSEVTPDOSKPEN 229
DB 181 NASEMDIQNVPLSEKIAELKEKIVLTNRKLSLMLKILSEVTPDOSKPEN 229

RESULT 3
AAM42009
ID AAM42009 standard; protein; 231 AA.
XX
XX AAM42009;
XX
DT 22-OCT-2001 (first entry)

XX DE Human polypeptide SEQ ID NO 6940.

XX KW Human; nototropic; immunosuppressant; cytostatic; gene therapy; cancer;

XX KW peripheral nervous system; neuropathy; central nervous system; CNS;

XX KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;

XX KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;

XX KW chemokine; thrombolytic; drug screening; arthritis; inflammation;

XX KW leukaemia.

XX OS Homo sapiens.

XX PN WO200153312-A1.

XX PD 26-JUL-2001.

XX PF 26-DEC-2000; 2000WO-US034263.

XX PR 23-DEC-1999; 99US-00471275.

XX PR 21-JAN-2000; 2000US-00488725.

XX PR 25-APR-2000; 2000US-0052317.

XX PR 20-JUN-2000; 2000US-00598042.

XX PR 19-JUL-2000; 2000US-00620312.

XX PR 03-AUG-2000; 2000US-00653450.

XX PR 14-SEP-2000; 2000US-00662191.

XX PR 19-OCT-2000; 2000US-00693036.

XX PR 29-NOV-2000; 2000US-00727344.

XX PA (HYSE-) HYSEQ INC.

XX PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;

XX PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;

XX PI Zhou P, Goodrich R, Drmanac RT;

XX DR WPI; 2001-442253/47.

XX DR N-PSDB; AAI61165.

XX PT Novel nucleic acids and polypeptides, useful for treating disorders such

XX PT as central nervous system injuries.

XX PS Example 2; SEQ ID NO 6940; 10078pp; English.

XX CC The invention relates to human nucleic acids (AAI57798-AAI61369) and the

XX CC encoded polypeptides (AAM38642-AAM42213) with nootropic,

XX CC immunosuppressant and cytostatic activity. The polynucleotides are useful

XX CC in gene therapy. A composition containing a polypeptide or polynucleotide

XX CC of the invention may be used to treat diseases of the peripheral nervous

XX CC system, such as peripheral nervous injuries, peripheral neuropathy and

XX CC localised neuropathies and central nervous system diseases, such as

XX CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic

XX CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the

XX CC utilisation of the activities such as: immune system suppression,

XX CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic

XX CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,

XX CC assays for receptor activity, arthritis and inflammation, leukaemias and

XX CC C.N.S disorders. Note: The sequence data for this patent did not form

XX CC part of the printed specification

XX SQ Sequence 231 AA;

Query Match 100.0%; Score 1198; DB 4; Length 231;

Best Local Similarity 100.0%; Pred. No. 9.8e-127;

Matches 229; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAAGPLHRRRCATPPRGDFCGGTGTAIDQASFTTSMEDTQVVGKSSPLGAGLGAEEP 60

DB 3 MAAGPLHRRRCATPPRGDFCGGTGTAIDQASFTTSMEDTQVVGKSSPLGAGLGAEEP 62

QY 61 AAGPQLPSWLOPERCAVFOCAQCHAVLADSVHLAWDLRSGLGAVFSRVNTNVVLEAPFL 120

DB 63 AAGPQLPSWLOPERCAVFOCAQCHAVLADSVHLAWDLRSGLGAVFSRVNTNVVLEAPFL 122

QY 121 VGIEGSLKGYSTNLLFCGSCGIPVGFHYSTHAALARGHFCLSDDKMWCVLLKTKAIV 180

DB 123 VGIEGSLKGYSTNLLFCGSCGIPVGFHYSTHAALARGHFCLSDDKMWCVLLKTKAIV 182

QY 181 NASEMDIQNVPLSEKIAELKEKIVLTHNRLSKMLKILSEVTPDQSKPEN 229

DB 183 NASEMDIQNVPLSEKIAELKEKIVLTHNRLSKMLKILSEVTPDQSKPEN 231

RESULT 4

ABG34855

ID ABG34855 standard; protein; 164 AA.

XX AC ABG34855;

XX DT 15-JUL-2002 (first entry)

XX DE Human cDNA 85PIB3 splice variant, open reading frame #3.

XX KW Human; cytostatic; 85PIB3; cancer; immunogen; chromosome 15q14.

XX OS Homo sapiens.

XX PN WO200218578-A2.

XX PD 07-MAR-2002.

XX PF 28-AUG-2001; 2001WO-US026838.

XX PR 28-AUG-2000; 2000US-0228432P.

XX PA (AGEN-) AGENSYS INC.

XX PI Raitano AB, Paris M, Hubert RS, Afar D, Ge W, Challita-Eid P;

XX PI Jakobovits A;

XX DR WPI; 2002-382963/41.

XX DR N-PSDB; ABK70504.

XX PT Composition for modulating the status of 85PIB3 protein or a molecule

XX PT comprising a substance e.g. antibody specific to, nucleic acid encoding,

XX PT or ribozyme of 85PIB3.

XX PS Example 38; Page 124; 201pp; English.

XX CC The invention relates to a composition comprising a substance that

XX CC modulate the status of 85PIB3, where the status of a cell expresses

XX CC 85PIB3 gene product is modulated. Also included are a composition

XX CC comprising a peptide region of 5 amino acids of the 85PIB3 protein, in

XX CC any whole number increment up to 229 that includes an aa position

XX CC selected from an aa position having a value greater than 0.5 in the

XX CC hydrophilicity profile, an aa position having a value less than 0.5 in

XX CC the hydrophobicity profile, an aa position having a value greater than

XX CC 0.5 in the percent accessible residue profile, an aa position having a

XX CC value greater than 0.5 in the average flexibility profile, or an aa

XX CC position having a value greater than 0.5 in the beta-turn profile; a

XX CC polynucleotide that encodes analogue peptide of 8, 9, 10 or 11 contiguous

XX CC residues of the 85PIB3 protein; a recombinant protein comprising the

XX CC antigen-binding region of a monoclonal antibody; a non-human transgenic

XX CC animal that produces an antibody that binds to the 85PIB3 protein; a

XX CC hybridoma that produces an antibody specific to the protein; a single chain

XX CC monoclonal antibody (Mab) that comprises the variable domains of the

XX CC heavy and monoclonal antibodies specific to the protein; a vector

XX CC comprising a polynucleotide that encodes the Mab; inhibiting growth of

XX CC cancer cells or treating a patient who bears cancer cells that expresses

XX CC the protein, by administering the protein, antibody, polynucleotide

XX CC encoding the protein, antisense polynucleotide to the polynucleotide,

XX CC ribozyme that cleaves the polynucleotide and T cells that specifically

XX CC recognize the protein; and generating a mammalian immune response

XX CC directed to the protein exposing cells of the mammal's immune system to

XX CC an immunogenic portion of the protein or polynucleotide. The composition,

XX CC which comprises an antibody specific to the protein, is useful for

XX CC delivering a cytotoxic agent to a cell that expresses the protein by

XX CC providing a cytotoxic agent conjugated to antibody and exposing the cell

CC to the antibody-agent conjugate. The methods are useful for inhibiting
 CC growth of cancer cells or treating a patient who bears cancer cells that
 CC expresses the protein, for generating a mammalian immune response
 CC directed to the protein, for detecting the presence of the protein or
 CC polynucleotide in a biological sample in a patient who has or who is
 CC suspected of having cancer, and for monitoring 85PIB3 in a biological
 CC sample from a patient who has or who is suspected of having cancer. The
 CC gene for 85PIB3 is located on human chromosome 15q14. The present
 CC sequence is the 85PIB3 protein

XX SQ Sequence 164 AA;

Query Match 11.6%; Score 138.5; DB 5; Length 164;
 Best Local Similarity 82.1%; Pred. No. 8.6e-07;
 Matches 32; Conservative 1; Mismatches 1; Indels 5; Gaps 1;

Qy 162 FCLSSDRMVCYLLKTKAIVNASEMDIONVPLSEKIAELK 200
 Db 2 FFLSS-----YLLTKAIVNASEMDIONVPLSEKIAEVK 35

RESULT 5

AAAM93668
 ID AM93668 standard; protein; 233 AA.

XX AC AM93668;

XX DT 06-NOV-2001 (first entry)

XX DE Human polypeptide, SEQ ID NO: 3550.

XX KW Human; full length cDNA; cDNA synthesis; oligo-capping.

XX OS Homo sapiens.

XX PN EP1130094-A2.

XX PD 05-SEP-2001.

XX PF 07-JUL-2000; 2000EP-00114089.

XX PR 08-JUL-1999; 99JP-00194486.

XX PR 11-JAN-2000; 2000JP-00118774.

XX PR 02-MAY-2000; 2000JP-00183765.

XX PA (HELI-) HELIX RES INST.

XX PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;

XX PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;

XX DR WPI; 2001-524255/58.

XX DR N-PSDB; AAK94603.

XX PT 830 Primers useful for synthesizing full length cDNA clones and their use

XX PS in genetic manipulation.

XX PS Claim 8; SEQ ID NO 3550; 1380pp + Sequence Listing; English.

XX The invention relates to primers for synthesizing full length cDNA
 CC clones, 830 cDNA molecules encoding a human protein have been isolated
 CC and nucleotide sequences of 5'- and 3'-ends of the cDNA molecules have
 CC been determined. Primers for synthesizing the full length cDNA are useful
 CC for clarifying the function of the protein encoded by the cDNA. The full
 CC length clones were obtained by construction of full length enriched cDNA
 CC libraries that were synthesised by the oligo-capping method. The primers
 CC enable the production of the full length cDNA easily without any special
 CC methods. The present sequence is a polypeptide encoded by a full length
 CC human cDNA of the invention. Note: The sequence data for this patent did
 CC not form part of the printed specification, but was obtained in CD-ROM
 CC format directly from EPO

XX SQ Sequence 233 AA;

Query Match 9.4%; Score 113; DB 4; Length 233;
 Best Local Similarity 27.1%; Pred. No. 0.0011;
 Matches 52; Conservative 30; Mismatches 88; Indels 22; Gaps 8;
 Qy 31 ASFTTSEWDTQVVKGSSPLGPAAGABEPAAGPQLPSWLQPERCAVFOCAQCHAVLADS 90
 Db 49 ASMWSSSEDAV-----ADMERAGL--EEENAAAE-----ERPLVFLCSGCCRPGLGDS 95
 Qy 91 VHLAWDLR--SLGAVFVSRTNNVYLRAPFLVGIKSLKSTYNNLLFCGSCGIPVGFHLY 149
 Db 96 --LSWVASQEDTNCILLRCVSCNVSDKEQKLSKREKNGCVLETLCAGCGLNLGYVYR 153
 Qy 150 STHAALAAALRGHPCILSSDKMVCYLL--KTKAIVNASEMDIONVPLSEKIAELKEKIVLTH 207
 Db 154 CTPKNLDYKRDLCFLSVEAIESYVLGSSEKQIV--SEDKLEFNL---ESRVEIEKSLTQME 209
 Qy 208 NRLKSLMKILSE 219
 Db 210 DVLKALQMKLWE 221

RESULT 6

AAAG93290

ID AG93290 standard; protein; 233 AA.

XX AC AG93290;

XX DT 13-SEP-2001 (first entry)

XX DE Human protein HP10650.

XX KW Human; gene therapy; tumour.

XX OS Homo sapiens.

XX PN WO200142302-A1.

XX PD 14-JUN-2001.

XX PF 06-DEC-2000; 2000WO-JP008631.

XX PR 06-DEC-1999; 99JP-00346863.

XX PR 06-DEC-1999; 99JP-00346864.

XX PR 08-FEB-2000; 2000JP-00031062.

XX PR 10-FEB-2000; 2000JP-00034090.

XX PR 10-FEB-2000; 2000JP-00034091.

XX PR 14-FEB-2000; 2000JP-00035829.

XX PR 14-FEB-2000; 2000JP-00035899.

XX PR 14-MAR-2000; 2000JP-00071161.

XX PR 30-MAY-2000; 2000JP-00160851.

XX PA (NISC-) JAPAN SCI & TECHNOLOGY CORP.

XX PI Kato S, Eguchi C, Saeki M;

XX DR WPI; 2001-381646/40.

XX DR N-PSDB; AAH68575.

XX PT Human protein originated from tumor cell line, applicable as drug,

XX PS reagent for studying intracellular protein networks and protein source

XX PS for drug screening, also encoded cDNA for gene diagnosis and gene

XX PT therapy.

XX PS Claim 1; Page 303-304; 471pp; Japanese.

XX The present sequence is a human protein. The human protein, preferably
 CC originated from tumour cell line, is applicable as a drug, a reagent for
 CC studying intracellular protein networks and a protein source for
 CC screening proteins for binding low molecular weight drugs. The human
 CC protein coding sequence is useful for gene diagnosis and gene therapy,
 CC expression vectors and transformant cells for detection of ligands and
 CC receptors

CC	present invention describes a method for identifying such proteins that
CC	can be used in the treatment of various diseases associated with altered
CC	mitochondrial function including diabetes mellitus, Huntington's disease
CC	osteoarthritis, Leber's hereditary optic neuropathy (LHON), mitochondria
CC	encephalopathy lactic acidosis and stroke (MELAS), myoclonic epilepsy
CC	ragged red fibre syndrome (MERRF) or cancer. Accordingly, these
CC	compositions have neuroprotective, nootropic, antidiabetic,
CC	anticonvulsant, antiarthritic, osteopathic, ophthalmological and
CC	cytostatic activities. This polypeptide sequence is a human heart
XX	mitochondrial protein of the invention.
SQ	Sequence 233 AA;
	Query Match 9.4%; Score 113; DB 7; Length 233;
	Best Local Similarity 27.1%; Pred. NO. 0.0011;
	Matches 52; Conservative 30; Mismatches 88; Indels 22; Gaps 14
QY	31 ASFTTSMEDWTGVKSSPLGAGLGAEPAPQLPSPWLOPRCAVFCQAACHAVLADS 90
DB	: : : : : :
DB	49 ASMSSMSSEDASV----ADMERAQL--EEENAAA-----ERPLVFLCSGCCRRRLGDS 95
QY	91 VHLAWDLGR-SLGAVVFGRVTNNVLAEPLVGIEGLKGSTYNLLFCGSCGIPVGPHLY 14
DB	: : : :
DB	96 --LSWVASQEDTNCLLRVCSCNVSVDKQKLKRKENGCVLETCCACCSNLGGVYR 15
QY	150 STHAALAAALRGHFCILSSDRKMVCYL--KTKAIVNASEMDIQNVPLSEKIAELKEKIVLTH 20
DB	
DB	154 CTPKNLDYKDFLCLSVEAIESYVLGSSEKQIV-SEDKLEFNL---ESRVEIEKSLTQME 205
QY	208 NRLKSIMKILSE 219
DB	:
DB	210 DVLXALOMKLWE 221
	RESULT 8
ID	ADL31517 standard; protein; 233 AA.
XX	ADL31517;
AC	ADL31517;
XX	20-MAY-2004 (first entry)
DT	Human protein encoded by a full length cDNA clone SeqID 3550.
DE	human; medicine; signal transduction; glycoprotein; transcription;
KW	oligo-capping method.
KW	Homo sapiens.
OS	EPI396543-A2.
PN	10-MAR-2004.
PD	07-JUL-2000; 2003EP-00025638.
Pf	08-JUL-1999; 99JP-00194486.
XX	11-JAN-2000; 2000JP-00118774.
PR	02-MAY-2000; 2000JP-00183865.
PR	07-JUL-2000; 2000EP-00114089.
XX	(REAS-) RES ASSOC BIOTECHNOLOGY.
PA	Ota T. Nishikawa T. Isogai T. Hayashi K, Iehli S, Kawai Y;
PI	Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
PI	WFI; 2004-204755/20.
XX	N-PSDE; ADL31516.
XX	New oligonucleotide primers (830 cDNAs) useful for synthesizing full
PT	length human cDNAs.
PT	Example 1; SEQ ID NO 3550; 1340pp; English.
XX	PS
XX	XX

CC This invention relates to a novel primers useful for synthesising full
CC length cDNA molecules that encode human proteins. Specifically, it refers
CC to secretory or membrane proteins that are potential therapeutic agents/
CC target molecules in the field of medicine, and in particular genes
CC encoding proteins that are associated with signal transduction.
CC glycoproteins and transcription. The present invention describes a method
CC for efficiently cloning a full length human cDNA from both the 5' and 3'
CC ends using the oligo-capping method. This polypeptide sequence is a full
CC length human protein of the invention.

XX SQ Sequence 233 AA;

Query Match 9.4%; Score 113; DB 8; Length 233;
Best Local Similarity 27.1%; Pred. No. 0.0011;
Matches 52; Conservative 30; Mismatches 88; Indels 22; Gaps 8;
QY 31 ASFTTSEWDTQVVKGSPLGPAGIGAEPAAGPOLPSWLQPERCAVQCAQCHAVLADS 90
DB 49 ASWSSMEDASV-----ADMERAGL--EEEAARAE-----ERPLVLCGCRPLGDS 95
QY 91 VHLAWDLNR-SLGAVFVSRTNNVLEAPFLVIGIEGLSGSTYNNLFCGSGCIPVGFHLY 149
DB 96 --LSWVASQEDTNCILLRCVSCNVSDKEQKLSKREKENGCVLETLCCAGCSNLNGVYR 153
QY 150 STHAALAAALRGHCLSSDKMVCYLL--KTKAIVNASEMDIQNVPLSEKIAELKEKIVLTH 207
DB 154 CTPKRLDYKRDLCFLSVEATESYVIGSSEKQIV-SEDKELFNL---ESRVEIEKSLTOME 209
QY 208 NRLKSLMKILSE 219
DB 210 DVLKALQKLWE 221

RESULT 9

ABB54167
ID ABB54167 standard; protein; 695 AA.

XX AC ABB54167;
XX DT 29-AUG-2003 (revised)
XX DT 16-MAY-2002 (first entry)
XX DE Lactococcus lactis protein yihC.
XX KW Biosynthesis; biodegradation; lactic bacterium; yogurt; cheese.

XX OS Lactococcus lactis; IL1403.

XX PN FR2807446-A1.

XX PD 12-OCT-2001.

XX PF 11-APR-2000; 2000FR-00004630.

XX PR 11-APR-2000; 2000FR-00004630.

XX PA (INRG) INRA INST NAT RECH AGRONOMIQUE.

XX PI Bolotine A, Sorokine A, Renault P, Ehrlich SD;

XX DR WPI; 2002-043418/06.

XX PT New nucleotide sequence useful in the identification or Lactococcus
XX PT lactis and related species.

XX PS Claim 6; SEQ ID NO 869; 2504pp; French.

XX CC The present invention is related to a Lactococcus lactis nucleotide
XX CC sequence (AB90521) and related proteins (ABB5300-ABB55621). The nucleic
XX CC acid sequence is useful in the detection and/or amplification of nucleic
XX CC acid sequence, particularly to identify Lactococcus lactis or related
XX CC species. The proteins of the invention are useful for the biosynthesis or
XX CC biodegradation of a composition of interest. The invention helps research

CC in lactic bacteria, particularly useful in the production of yogurt and
CC cheese. Note: The sequence data for this patent is based on equivalent
CC patent WO200177334 (published 18-OCT-2001) which is available in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences. (Updated on 29-AUG-2003 to
CC standardise OS field)

XX SQ Sequence 695 AA;

Query Match 7.7%; Score 92; DB 5; Length 695;
Best Local Similarity 25.3%; Pred. No. 1.5;
Matches 38; Conservative 28; Mismatches 50; Indels 34; Gaps 7;
QY 101 LGAVFVSRTNNVLEAPFLVIGIEGLSGSTYNNLFCGSGCIPVGFHLYSTHAALALRG 160
DB 101 LGTIIF-----FYSGTFFSGAGELSKRPAMMLITMGITVAY-AYSVIATIMSLNG 153
QY 161 HF-----CLSSDKMVCYLLKTKAIVNASE--MDI-----QNVPLSE-K 195
DB 154 HMGNFVPELATLIVIMLIGHLIEMKAIMGAGDALDASLVPKKAHLKSGKDVSELK 213
QY 196 IAEKKEKIVLTHNRLKSLMKILSEVTPDQS 225
DB 214 VGDJ--LLVKEKEKIPADGLILSEALVDES 241

RESULT 10

ADS29356
ID ADS29356 standard; protein; 695 AA.

XX AC ADS29356;

XX DT 02-DEC-2004 (first entry)

XX DE Bacterial polypeptide #18389.

XX KW Recombinant DNA construct; transformed plant; improved plant property;
XX KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
XX KW pathogen tolerance; pest tolerance; plant disease resistance;
XX KW cell cycle pathway modification; plant growth regulator;
XX KW homologous recombination; seed oil yield; protein yield; carbohydrate;
XX KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
XX KW bacterial polypeptide.

XX OS Bacteria.

XX PN US2003233675-A1.

XX PD 18-DEC-2003.

XX PF 20-FEB-2003; 2003US-00369493.

XX PR 21-FEB-2002; 2002US-0360039P.

XX PA (CAOY/) CAO Y.

XX PA (HINK/) HINKLE G J.

XX PA (SLAT/) SLATER S C.

XX PA (CHEN/) CHEN X.

XX PA (GOLD/) GOLDMAN B S.

XX PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;

XX DR WPI; 2004-061375/06.

XX PT New recombinant DNA construct comprising a promoter positioned to provide
XX PT for expression of a polynucleotide encoding a polypeptide from a
XX PT microbial source, useful for producing plants with improved properties.

XX PS Claim 1; SEQ ID NO 18389; 122pp; English.

XX CC The invention relates to a recombinant DNA construct comprising a
XX CC promoter functional in a plant cell, where the promoter is positioned to
XX CC provide for expression of a polynucleotide encoding a polypeptide from a

PI Legace RE, Spiro BA, Stewart EA, Wingrove J, Vitt UA, Kirton ES;
PI Xu Y, Kwong M, Policky JL, Hurwitz BL, Ma Y, Jackson JL, Gietzen D;
PI Patury S, Shi X, Suarez CJ;
XX WPI; 2004-329368/30.
DR N-PSDB; ACN42006.
XX
XX New diagnostic and therapeutic polynucleotides and polypeptides, useful
PT in diagnosing a condition, disease or disorder associated with human
PT molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or
PT in gene mapping.
XX
PS Claim 27; Page; 190pp; English.
XX
XX The invention relates to novel diagnostic and therapeutic polynucleotides
CC selected from one of the 2722 sequences defined in the specification. A
CC polynucleotide of the invention may have a use in gene therapy. The human
CC diagnostic and therapeutic polynucleotides (dithp) or polypeptides may be
CC used to diagnose a particular condition, disease or disorder associated
CC with human molecules, e.g. cell proliferative disorders,
CC autoimmune/inflammatory disorder, developmental disorder, endocrine
CC disorder, neurological disorders, gastrointestinal disorders, or
CC infections caused by virus, bacteria, fungi or parasite. The dithp
CC molecules may also be used in genetic mapping, in identifying individuals
CC from minute biological samples, in detecting single nucleotide
CC polymorphisms, as molecular weight markers, and for somatic or germline
CC gene therapy. The present sequence represents a dithp protein of the
CC invention. Note: The sequence data for this patent is not represented in
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at www.wipo.int/pct/en/sequences/listing.htm
XX
SQ Sequence 572 AA;

Query Match 7.4%; Score 89; DB 8; Length 572;
Best Local Similarity 23.0%; Pred. No. 2.4;
Matches 52; Conservative 26; Mismatches 74; Indels 74; Gaps 11;

QY 8 HRSRCATPPRGDFC-----GCTERAIDQA-----SFTSMEDTQVVGSSPL----- 50
DB 342 YRQQAQPPH---CPAPEGEPGAPQALGDAPTSVSLTTAVQDVVCPQSGSHALCTCCFQ 398
QY 51 ---GPAGLGAEEPAAGPQLPSWLQPERCAVFOCAQCHAVLADSVHLAWDLR----- 99
DB 399 PMPDRAERQDPRVAPQ-----QCAVC---LQPFCHLYWGCTRTGCGYCLA 442
QY 100 -----SLGAVVFSRVNNVLEAPFLVIGEGSLKGSTY-NLLFCGSGIPVGFHLVSTHA 153
DB 443 PFCELNLGDKCLDGLVNNNSYESDILKNYLAT-RGLTWKNML-----TES 486
QY 154 ALAALRGHFCLS-----SDKMVCYLLKTKAIVNASEMDIONVPLSE 194
DB 487 LVALQRGVFLLSDYRVVTGDTVLCYCGLRSFRELITYQYQONIPASE 532

RESULT 13
AAB93182
ID AAB93182 standard; protein; 623 AA.
XX
XX AAB93182;
XX
XX 26-JUN-2001 (first entry)
XX
XX Human protein sequence SEQ ID NO:12128.
XX
XX Human; primer; detection; diagnosis; antisense therapy; gene therapy.
XX
XX Homo sapiens.
XX
XX EP1074617-A2.
XX
XX 07-FEB-2001.
XX
XX 28-JUL-2000; 2000EP-00116126.

XX 29-JUL-1999; 99JP-00248036.
PR 27-AUG-1999; 99JP-00300253.
PR 11-JAN-2000; 2000JP-00118776.
PR 02-MAY-2000; 2000JP-00183767.
PR 09-JUN-2000; 2000JP-00241899.
XX
XX (HELI-) HELIX RES INST.
XX
XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX WPI; 2001-318749/34.
XX
XX Primer sets for synthesizing polynucleotides, particularly the 5602 full-
PT length cDNAs defined in the specification, and for the detection and/or
PT diagnosis of the abnormality of the proteins encoded by the full-length
PT cDNAs.
XX
PS Claim 8; SEQ ID NO 12128; 2537pp + Sequence Listing; English.
XX
XX The present invention describes primer sets for synthesizing 5602 full-
CC length cDNAs defined in the specification. Where a primer set comprises:
CC (a) an oligo-dT primer and an oligonucleotide complementary to the
CC complementary strand of a polynucleotide which comprises one of the 5602
CC nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in the
CC specification. The primer sets can be used in antisense therapy and in
CC gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
CC oligonucleotides, all of which are used in the exemplification of the
CC present invention
XX
SQ Sequence 623 AA;

Query Match 7.4%; Score 89; DB 4; Length 623;
Best Local Similarity 23.0%; Pred. No. 2.7;
Matches 52; Conservative 26; Mismatches 74; Indels 74; Gaps 11;

QY 8 HRSRCATPPRGDFC-----GCTERAIDQA-----SFTSMEDTQVVGSSPL----- 50
DB 393 YRQQAQPPH---CPAPEGEPGAPQALGDAPTSVSLTTAVQDVVCPQSGSHALCTCCFQ 449
QY 51 ---GPAGLGAEEPAAGPQLPSWLQPERCAVFOCAQCHAVLADSVHLAWDLR----- 99
DB 450 PMPDRAERQDPRVAPQ-----QCAVC---LQPFCHLYWGCTRTGCGYCLA 493
QY 100 -----SLGAVVFSRVNNVLEAPFLVIGEGSLKGSTY-NLLFCGSGIPVGFHLVSTHA 153
DB 494 PFCELNLGDKCLDGLVNNNSYESDILKNYLAT-RGLTWKNML-----TES 537
QY 154 ALAALRGHFCLS-----SDKMVCYLLKTKAIVNASEMDIONVPLSE 194
DB 538 LVALQRGVFLLSDYRVVTGDTVLCYCGLRSFRELITYQYQONIPASE 583

RESULT 14
AAB97233
ID AAB97233 standard; protein; 623 AA.
XX
XX AAB97233;
XX


```
DT 27-JUN-2002 (first entry)
XX
XX
DE Novel human protein SEQ ID NO: 501.
XX
XX Human; antianaemic; vulnerary; antiinflammatory; immunomodulator;
XX antiinfertility; cerebroprotective; cytostatic; rheumatic; gene therapy;
XX neuroprotective; antiparkinsonian; protein therapy; EST;
XX expressed sequence tag.
XX
XX Homo sapiens.
XX
XX WO200222660-A2.
XX
XX 21-MAR-2002.
XX
XX 10-SEP-2001; 2001WO-US026015.
XX
XX 11-SEP-2000; 2000US-00659671.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;
XX Xue AJ, Yang Y, Wehrman T, Drmanac RT;
XX
XX WPI; 2002-292408/33.
XX
XX N-PSDB; ABN32419.
XX
XX An isolated polynucleotide for treating diseases associated with its
XX encoded polypeptide such as cancer and multiple sclerosis.
XX
XX Example 2; SEQ ID NO 501; 509pp; English.
XX
XX The present invention provides the protein and coding sequences of 444
XX novel human proteins. These were isolated from expressed sequences tags
XX (ESTs). They can be used to stimulate cell growth, to regulate
XX haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth
XX e.g. in burn treatment, to regulate the immune system e.g. to treat
XX multiple sclerosis, to regulate activin or inhibin e.g. to treat
XX infertility, to regulate haemostasis or thrombolysis e.g. to treat stroke
XX and cancer, to screen for drugs, to treat inflammatory conditions e.g.
XX rheumatoid arthritis, and to treat nervous system disorders e.g.
XX Parkinson's disease. The present sequence is a protein of the invention
XX
XX Sequence 623 AA;
XX
XX Query Match 7.4%; Score 89; DB 5; Length 623;
XX Best Local Similarity 23.0%; Pred. No. 2.7;
XX Matches 52; Conservative 26; Mismatches 74; Indels 74; Gaps 11;
XX
XX QY 8 HRSRCATPPRGDFC-----GGTERAIDQA-----SFTTSMEDTQVVKGSPL----- 50
XX Db 393 YRQAAQPPH---CPAPEGEPGAPQALGDAPPTSVSLTTAVQDYVCPLOGSHALCTCCFQ 449
XX
XX QY 51 ---GPAGLGAEEPAAGPOLPSWLPQERCAVFOCAQCHAVLADSVHLAWDLR----- 99
XX Db 450 PMPDRAEREQDPVAPQ-----QCAVC---LQPFCHLYWCCTRTGCGVGLA 493
XX
XX QY 100 -----SLGAVVFSRVNTNNVLEAPFLVIGESLKGSTY-NLLFCGSCGIPVGFHLYSTHA 153
XX Db 494 PFCELNLGDKCLDGVLLNNNSYESDILKNYLAT-RGLTWKNML-----TES 537
XX
XX QY 154 ALAALRGHFCLS-----SPKWCYLLKTKAIVNASSEMDIQNVPLSE 194
XX Db 538 LVALQGVFLLSDYRVTGDTVLVCCGLRSFRELTYQOONIPASE 583
XX
XX RESULT 15
XX AAB93168
XX ID AAB93168 standard; protein; 652 AA.
XX
XX AAB93168;
XX
XX 26-JUN-2001 (first entry)
XX
XX Human protein sequence SEQ ID NO:12100.
XX
XX Human; primer; detection; diagnosis; antisense therapy; gene therapy.
XX
XX Homo sapiens.
XX
XX EP1074617-A2.
XX
XX 07-FEB-2001.
XX
XX 28-JUL-2000; 2000EP-00116126.
XX
XX 29-JUL-1999; 99JP-00248036.
XX
XX 27-AUG-1999; 99JP-00300253.
XX
XX 11-JAN-2000; 2000JP-00118776.
XX
XX 02-MAY-2000; 2000JP-00183767.
XX
XX 09-JUN-2000; 2000JP-00241899.
XX
XX (HELI-) HELIX RES INST.
XX
XX Ota T, Isogai T, Nishikawa T, Hayaashi K, Saito K, Yamamoto J;
XX Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
XX WPI; 2001-318749/34.
XX
XX Primer sets for synthesizing polynucleotides, particularly the 5602 full-
XX length cDNAs defined in the specification, and for the detection and/or
XX diagnosis of the abnormality of the proteins encoded by the full-length
XX cDNAs.
XX
XX Claim 8; SEQ ID NO 12100; 2537pp + Sequence Listing; English.
XX
XX The present invention describes primer sets for synthesising 5602 full-
XX length cDNAs defined in the specification. Where a primer set comprises:
XX (a) an oligo-dT primer and an oligonucleotide complementary to the
XX complementary strand of a polynucleotide which comprises one of the 5602
XX nucleotide sequences defined in the specification, where the
XX oligonucleotide comprises at least 15 nucleotides; or (b) a combination
XX of an oligonucleotide comprising a sequence complementary to the
XX complementary strand of a polynucleotide which comprises a 5'-end
XX sequence and an oligonucleotide comprising a sequence complementary to a
XX polynucleotide which comprises a 3'-end sequence, where the
XX oligonucleotide comprises at least 15 nucleotides and the combination of
XX the 5'-end sequence/3'-end sequence is selected from those defined in the
XX specification. The primer sets can be used in antisense therapy and in
XX gene therapy. The primers are useful for synthesising polynucleotides,
XX particularly full-length cDNAs. The primers are also useful for the
XX detection and/or diagnosis of the abnormality of the proteins encoded by
XX the full-length cDNAs. The primers allow obtaining of the full-length
XX cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
XX AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
XX represent human amino acid sequences; and AAH13629 to AAH13632 represent
XX oligonucleotides, all of which are used in the exemplification of the
XX present invention
XX
XX Sequence 652 AA;
XX
XX Query Match 7.4%; Score 89; DB 4; Length 652;
XX Best Local Similarity 23.0%; Pred. No. 2.9;
XX Matches 52; Conservative 26; Mismatches 74; Indels 74; Gaps 11;
XX
XX QY 8 HRSRCATPPRGDFC-----GGTERAIDQA-----SFTTSMEDTQVVKGSPL----- 50
XX Db 422 YRQAAQPPH---CPAPEGEPGAPQALGDAPPTSVSLTTAVQDYVCPLOGSHALCTCCFQ 478
XX
XX QY 51 ---GPAGLGAEEPAAGPOLPSWLPQERCAVFOCAQCHAVLADSVHLAWDLR----- 99
XX Db 479 PMPDRAEREQDPVAPQ-----QCAVC---LQPFCHLYWCCTRTGCGVGLA 522
XX
XX QY 100 -----SLGAVVFSRVNTNNVLEAPFLVIGESLKGSTY-NLLFCGSCGIPVGFHLYSTHA 153
XX Db 523 PFCELNLGDKCLDGVLLNNNSYESDILKNYLAT-RGLTWKNML-----TES 566
```

Qy 154 ALAALRGHFCLS-----SDXWVCYLLKTKAIVNASEMDIQNVPLSE 194
Db 567 LVALORGVELLSDYRVGTDLVCYCCGLRSFRELTYQYRQNPASE 612

Search completed: March 25, 2006, 07:59:14
Job time : 191 secs

Copyright (c) 1993 - 2006 Bioceleration Ltd.
GENCODE version 5.1.7

Result No.	Query %			DB	ID	Description
	Score	Match	Length			
1	92	7.7	695	2	C86731	copper-potassium t
2	90.5	7.6	597	2	D71293	probable phosphori
3	89	7.4	306	2	T46399	hypothetical prote
4	88	7.3	250	2	A43623	kanamycin kinase (
5	86.5	7.2	669	2	A97229	NAD-dependent DNA
6	85	7.1	2194	1	QJ1977	glutamate synthase
7	84.5	7.1	361	2	AE3074	ferrichrome bindin
8	83	6.9	663	2	D97047	DNA ligase (NAD de
9	82.5	6.9	143	2	B69099	formate hydrogenly
10	82.5	6.9	493	2	A85433	sugar transporter
11	82	6.8	967	2	F87678	DNA polymerase I (
12	81.5	6.8	329	2	D90404	transport protein,
13	81.5	6.8	457	2	H85767	glucuronide permea
14	81.5	6.8	457	2	C90919	glucuronide permea
15	81.5	6.8	457	2	B64918	glucuronide permea
16	81.5	6.8	640	2	T41977	hypothetical prote
17	81.5	6.8	662	2	T44221	probable capsid pr
18	81	6.8	1418	2	A07664	hypothetical prote
19	80.5	6.7	766	2	B85440	receptor kinase-II
20	80.5	6.7	1711	1	A55148	protein-tyrosine-p
21	80	6.7	469	2	S55167	IME2-dependent sig
22	79	6.6	386	2	H90789	probable aminometh
23	79	6.6	386	2	C85650	probable aminometh
24	79	6.6	437	2	C86823	GTP-binding protei
25	78.5	6.6	188	2	F83816	late competence op
26	78.5	6.6	578	2	F86484	probable hydroxyme
27	78.5	6.6	697	1	Q8BEK3	Uil104 protein - hu
28	78	6.5	351	2	H82098	UDP-3-O-3-hydroxym
29	77.5	6.5	662	2	T44036	hypothetical prote

A;Residues: 1-597 <COL>
A;Cross-references: UNIPROT:O83693; UNIPARC:UPI00000D3290; GB:AE001243; GB:AE000520; NID
A;Experimental source: strain Nichols
C;Genetics:
A;Gene: TP0695

Query Match 7.6%; Score 90.5; DB 2; Length 597;
Best Local Similarity 28.0%; Pred. No. 2.2;
Matches 51; Conservative 19; Mismatches 47; Indels 65; Gaps 12;
QY 12 CATPRGDFCGTERRAID-----QASFT-----TSMEDTQVVKGSPLGPA 53
DB 90 CALP--GHRLEATKATKTRACFTTRALRCRFTFLPDSPAWDT-----PPGHA 140
QY 54 GLGAEEPAAGPOLPSWLOP-ERCAVFOC--AQCHAVLADSVHLAWDLRSRGAVVFSRVT 110
DB 141 RLCSHLHAGLSFPLVVKPTDNMGARGCTLAQCKDTLINACAVARQFSRS-----190
QY 111 NNVLAEAPLVGIEGSLKSTYNLLFCGSCGIPVGFHLYSTHAALALRGHFCILSSDKWV 170
DB 191 GRVIEE-BFIVGREFSLEG---LIFDGT-----LYVT--ALA-----DRHI 225
QY 171 CY 172
DB 226 CF 227

RESULT 3
T46399
hypothetical protein DKFp434N2420.1 - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 09-Jul-2004
C;Accession: T46399
R;Ottenwaelder, B.; Obermaier, B.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, January 2000
A;Reference number: Z33031
A;Accession: T46399
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-306 <AAA>
A;Cross-references: UNIPROT:Q96BP1; UNIPARC:UPI000006D779; EMBL:AL137561
A;Experimental source: adult testis; clone DKFp434N2420
C;Genetics:
A;Note: DKFp434N2420.1

Query Match 7.4%; Score 89; DB 2; Length 306;
Best Local Similarity 23.0%; Pred. No. 1.3;
Matches 52; Conservative 26; Mismatches 74; Indels 74; Gaps 11;
QY 8 HRSRCATPPRGDFC-----CGTERAIDQA-----SFTTSMEDTQVVKGSPL-----50
DB 76 YRQAQAPPH---CPAPEGEPCAPQALGDAPSTSVSLTAVQDVYCPQLQGSHALCTCCFQ 132
QY 51 ---GPAGLGAEPAAGPOLPSWLOPERCAVFOCAQCHAVLADSVHLAWDLRS-----99
DB 133 PMPDRAERQDPVAFQ-----QCAVC---LQPFCHLYWGCTRTCCYCGCLA 176
QY 100 -----SLGAVVFSRVTNNVLAEAPLVGIEGSLKSTY-NLLFCGSCGIPVGFHLYSTHA 153
DB 177 PFCEINLGDKLDGVLNNNSVESDILKNYLAT-RGLTWKNML-----TES 220
QY 154 ALAALRGHFCLS-----SDKMVCYLLTKTAIVNASEMDIONVPLSE 194
DB 221 LVALQRGVLLSDYRVGTDTLVLCYCCGLRSFREUTYQYRQNPAPASE 266

RESULT 4
A43623
kanamycin kinase (EC 2.7.1.95) - Campylobacter jejuni
C;Species: Campylobacter jejuni
C;Date: 11-Feb-1993 #sequence_revision 11-Feb-1993 #text_change 31-Dec-2004
C;Accession: A43623
R;Tenover, F.C.; Gilbert, T.; O'Hara, P.

Plasmid 22, 52-58, 1989
A;Title: Nucleotide sequence of a novel kanamycin resistance gene, aphA-7, from Campylobacter
A;Reference number: A43623; MUID:89387451; PMID:2550983
A;Accession: A43623
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-250 <TEN>
A;Cross-references: UNIPROT:P14508; UNIPARC:UPI000012DBEC; GB:M29953; GB:J03316; NID:91
C;Superfamily: aminoglycoside 3'-phosphotransferase (kanamycin kinase)
C;Keywords: phosphotransferase

Query Match 7.3%; Score 88; DB 2; Length 250;
Best Local Similarity 28.7%; Pred. No. 1.3;
Matches 33; Conservative 18; Mismatches 46; Indels 18; Gaps 5;
QY 105 VFSRVTNNVLAEAPLVGIEGSLKSTYNLLFCGSCGIPVGFHLYSTHAALALRG-HFC 163
DB 40 IFSKTTYSVREAEAMMWLSDKLKVPDV-----IEYGVREHSEYLIWSELRGKHID 90
QY 164 LSSDKMVCYLLTKTAIVNA-----SEMDIONVPLSEKI-AELKEKIVLTHNLKSL 213
DB 91 CFIDHPKIYI---ECLVNLHQLQAIDIRNCFPSKIDVRLKELKYLDDNRIADI 142

RESULT 5
A97229
NAD-dependent DNA ligase [imported] - Clostridium acetobutylicum
C;Species: Clostridium acetobutylicum
C;Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 31-Dec-2004
C;Accession: A97229
R;Nolling, J.; Breston, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee
; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A;Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium C1
A;Reference number: A96900; MUID:21359325; PMID:21359325
A;Accession: A97229
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-669 <KUR>
A;Cross-references: UNIPROT:Q97FQ5; UNIPARC:UPI000000C5D7; GB:AE001437; PIDN:AAK80620.1
A;Experimental source: Clostridium acetobutylicum ATCC824
C;Genetics:
A;Gene: CAC2673
C;Superfamily: DNA ligase (NAD), Liga type

Query Match 7.2%; Score 86.5; DB 2; Length 669;
Best Local Similarity 27.4%; Pred. No. 6.2;
Matches 37; Conservative 22; Mismatches 45; Indels 31; Gaps 8;
QY 96 DLSR---SLGNVFSRVTNNVLAEAPLVGIEGSLKSTYNL-----LFCGSCGIPVGF 146
DB 366 DIGRKKVKVIGSRVFRSRNDVI---PEIMGVTEGETNEIEAPTICPGSEIVKEGV 422
QY 147 HLYSTHAALALRGHFCILSSDKMVCYLLTKTAIVNASEMDIONVP-LSEKIAE-LKEKIV 204
DB 423 HL-----FC--ENTLSCKPQWKSIVHFASREAMNIEGFSEKTAQLFEK--465
QY 205 LTHNLKSLMKILSE 219
DB 466 LNIKISIDLXRTKE 480

RESULT 6
JQ1977
glutamate synthase (NADH2) (EC 1.4.1.14) precursor [validated] - alfalfa
C;Species: Medicago sativa (alfalfa)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 05-Oct-2004
C;Accession: JQ1977; PQ0551
R;Gregerson, R.G.; Miller, S.S.; Tuary, S.N.; Gantt, J.S.; Vance, C.P.
Plant Cell 5, 215-226, 1993
A;Title: Molecular characterization of NADH-dependent glutamate synthase from alfalfa r
A;Reference number: JQ1977; MUID:93200806; PMID:8453303
A;Accession: JQ1977

RESULT 8

D97047
DNA ligase (NAD dependent), Lga [imported] - Clostridium acetobutylicum
C/Species: Clostridium acetobutylicum
C/Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 31-Dec-2004
C/Accession: D97047
R./Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,
J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A./Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium C
A./Reference number: A96900; MUID:21359325; PMID:21359325
A./Accession: D97047
A./Status: preliminary
A./Molecule type: DNA
A./Residues: 1-663 <KUR>
A./Cross-references: UNIPROT:Q97JS8; UNIPARC:UPI00000CA10F; GB:AE001437; PIDN:AAK79167
A./Experimental source: Clostridium acetobutylicum ATCC824
C./Genetics:
A./Gene: CAC1195
C./Superfamily: DNA ligase (NAD), Lga type

Query Match 6.9%; Score 83; DB 2; Length 663;
Best Local Similarity 28.6%; Pred. No. 13;
Matches 36; Conservative 18; Mismatches 44; Indels 28; Gaps 7;

QY 102 GAVVPSRVNNVVLEAPFLVGIEGSLKGSTYNL-----LFCGSCGIPGVFHLYSTHAAL 155
||| :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 368 GARVLRSNDVI---PEIMGVTETEGETKEIEAPTICPYCGSRIVKGVHL----- 417
||| :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|

QY 156 AALRGHGLSSDKWCVLLTKTKAIVNASMDIQNVPLSEKIAELKEKIVLTNRRLKS 213
||| :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 418 -----FC--ENTLSCKPQWKVSI VHFASRKAMNIEGFSEKTAQLFEK--LNISIDSL 467
||| :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|

QY 214 MKILSE 219
:||:|
Db 468 YRITKE 473

RESULT 9

B69099
formate hydrogenlyase, iron-sulfur subunit 2 - Methanobacterium thermoautotrophicum (S
C/Species: Methanobacterium thermoautotrophicum
C/Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C/Accession: B69099
R./Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.
i. Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwan,
J. S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
J. Bacteriol. 179, 7135-7155, 1997
A./Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: fun
A./Reference number: A69000; MUID:98037514; PMID:9371463
A./Accession: B69099
A./Status: preliminary; nucleic acid sequence not shown; translation not shown
A./Molecule type: DNA
A./Residues: 1-143 <MTH>
A./Cross-references: UNIPROT:O27769; UNIPARC:UPI0000066781; GB:AE000929; GB:AE000666; N
A./Experimental source: strain Delta H
C./Genetics:
A./Gene: MTH1736
C./Superfamily: nrfC protein; ferredoxin 2[4Fe-4S] homology
F; 62-118/Domain: ferredoxin 2[4Fe-4S] homology <PER2>

Query Match 6.9%; Score 82.5; DB 2; Length 143;
Best Local Similarity 23.0%; Pred. No. 2.1;
Matches 35; Conservative 26; Mismatches 66; Indels 25; Gaps 6;

QY 72 PERCANVQCQAACHVLAADSVHLNDLSLGAVVF-----SRVNNVVLEAPFLVGIEG 125
||| :|:~|:|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|
Db 8 PELCD--ECMKCRIPCNAIRVID----GVPFCMHCSPERAPCLNICPEDAIVEVDG 60
||| :|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|

QY 126 SLKGSNTLLFFCSC--GIPIPVFHLYSTHAALAARGHGFLSSDKWCVLLTKTKAIVNAS 183
||| :|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|

Db 61 AVVILEDRICGCLCRDACPVG--AITLNERGVAVKDLICIDRDKPLCMVCPKGALES 118
Qy 184 EMDIQNVPLSEKIAELKKEKIVLTHNRLKSLMK 215
Db 119 SEDM-----MAAKRDKIAGELKRLKMLK 142

RESULT 10
A5433
sugar transporter like protein [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C:Accession: A85433
R:anonymus, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring
Nature 402, 769-777, 1999
A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
A:Reference number: A85001; MUID:20083488; PMID:10617198
A:Accession: A85433
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-493 <STO>
A:Cross-references: UNIPROT:O23213; UNIPARC:UPI000009F22F; GB:NC_001268; NID:G7270615; F
C:Genetics:
A:Map position: 4
A:Superfamily: glucose transport protein

Query Match 6.9%; Score 82.5; DB 2; Length 493;
Best Local Similarity 24.9%; Pred. No. 10;
Matches 53; Conservative 32; Mismatches 59; Indels 69; Gaps 13;

Qy 79 QCAQCHAVLADSVHLAWLSRLGAVVP---SRVTNNVLEAPFLVGI-----EGSLK- 128
Db 19 QCAIVASIV--SIIPGYDTGVMSGAMVFIEDLTKTDQIEV--LTGTLNLCALVGSLLA 74
Qy 129 -----GSTYNLLFCGSC--GIPVGFHL-----YSTH 152
Db 75 GRTSDIGRRYTVIVLASILFMLGSLMGWGNYPVLLSGRTAGLVGFGFALMVAPVSAE 134
Qy 153 AALAALRG-----HFLSSDRWCYLLK---TKAIVNAS---EMDIQNV---LSEKIA 197
Db 135 IATASHRGLLASPLHLCISIGILLGYVNYFFSKLPMHIGWRLMGLIAAAPSILVAFGIL 194
Qy 198 ELKE--KIVLTHNRLKSLMKILSEV--TPDQSK 226
Db 195 KMPSPRWLIMQGRLEKGEKILELVSNSPESAE 227

RESULT 11
F87678
DNA polymerase I [imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C:Accession: F87678
R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: A87249; MUID:21173698; PMID:11259647
A:Accession: F87678
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-967 <STO>
A:Cross-references: UNIPROT:Q9A2U2; UNIPARC:UPI00000C7AB1; GB:AE005673; NID:GI3425184; F
C:Genetics:
A:Gene: CC3464
C:Superfamily: DNA-directed DNA polymerase I

Query Match 6.8%; Score 82; DB 2; Length 967;
Best Local Similarity 24.1%; Pred. No. 26;
Matches 51; Conservative 27; Mismatches 72; Indels 62; Gaps 11;

Qy 12 CATPRGDFCGGTERAIDQ---ASFTTSMEDWT---QVVKSSPLGPRAGLGAEBEPAAGPQ 65
Db 271 CDTPLPQPLDALTVREPDKALAAFLQMEFRSLARRVGDGSAATPGTL--DRPAAPPK 328
Qy 66 LP-----SWL-----OPERC--AVFOCAQCHAVLADSVHLAWLSRLGAVVFSR 108
Db 329 APVSVSYNGAAARAAHPVPVKIDHAYACRDLATLKAWAKATD-----KGLVAFDT 384
Qy 109 VTNNVLEAPFLVGIKSLKSTYNLLFCGSCGIPVGFHLYSTHAALALRGHFLCSSDK 168
Db 385 ETD-----ALSSAT-----AGLCGV-----SLAIPCEACVIP-- 412
Qy 169 MCVYLLKTKAIVNASEMDIQNVPLSEKIAELK 200
Db 413 -ISHCEKADGLAFAPADIEQIPLADVIATLK 443

RESULT 12
D90404
transport protein, probable [imported] - Sulfolobus solfataricus
C:Species: Sulfolobus solfataricus
C:Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 09-Jul-2004
C:Accession: D90404
R:She, O.; Singh, R.K.; Confalonieri, P.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Char
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-NGOC, H.P.; Redder,
arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.
submitted to GenBank, April 2001
A:Description: Sulfolobus solfataricus complete genome.
A:Reference number: A99139
A:Accession: D90404
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-329 <KUR>
A:Cross-references: UNIPROT:Q97W97; UNIPARC:UPI00000646FA; GB:AE006641; NID:GI3815639;
C:Genetics:
A:Gene: SS02338

Query Match 6.8%; Score 81.5; DB 2; Length 329;
Best Local Similarity 22.2%; Pred. No. 7.4;
Matches 40; Conservative 30; Mismatches 47; Indels 63; Gaps 7;

Qy 87 LADSVHLA-WP-----LSRSLGAVVFSRVTNNV-----LE 116
Db 33 LSESHLAYMEVFAIVLALPFLGRIIGSFYQVFNKSVISYCFPLGLVILQNLGALIF 92
Qy 117 APFLVG-----IEGSLKSTYNLLFCGSCGIPVGFHLYSTHAALALRGHFLCSS 166
Db 93 VRFLVGVIIFGLLTSYAVESAVKSGRNVLVGTTAGWPIGWVI----- 134
Qy 167 DKWVCY-LLKTKAIVNASEMDIQNVPLSE---KIAELKEKIVLTHNRLKSLMKILSEVTP 222
Db 135 -SYVAYVLLKNWNVINISGILIMLALFELNGKEFGERSKISVSPRLTSLIYVSALTTP 193

RESULT 13
H85767
glucuronide permease [imported] - Escherichia coli (strain O157:H7, substrain EDL933)
C:Species: Escherichia coli
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C:Accession: H85767
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhe
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimailanta, E.; Potamouais, K.; Apodaca
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: H85767
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-457 <STO>
A:Cross-references: UNIPROT:Q8X673; UNIPARC:UPI00000DDBE5; GB:AE005174; NID:GI2515601;
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: uidB

C;Superfamily: melibiose carrier protein

Query Match 6.8%; Score 81.5; DB 2; Length 457;

Best Local Similarity 25.2%; Pred. No. 11;

Matches 37; Conservative 30; Mismatches 53; Indels 27; Gaps 8;

QY 95 WDLRSILGAVFVSRTNNV-----VLEA-----PFLVGIEGSLKSTYNLL-FCGSCG 141
Db 317 WSLPVALVALATASIGQGVMTVMWALEADTVVEGYLTGVR--IEGLTYSLSFSTRKCG 374
QY 142 IPVGPHLYSTHAALAAALRGHFC--LSSDKMVCYLLKTKAIVNASEMDIQNV-----PLSE 194
Db 375 QAIG--GSIPAFILGLSGYIANQVTPVIMGIRTSIALVPCGFMLLAFVLIWFYPLTD 431
QY 195 KIAELKEKIVLTHNRKLSMKILSEVT 221
Db 432 K--KPKFIVEIDNRKKVQQQLISDIT 456

RESULT 14

C90919 glucuronide permease [imported] - Escherichia coli (strain O157:H7, substrain RIMD 05099

C;Species: Escherichia coli

C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004

C;Accession: C90919

R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.

gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shingawa, H.

DNA Res. 8, 11-22, 2001

A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and geno

A;Reference number: A99629; MUID:21156231; PMID:11258796

A;Accession: C90919

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-457 <HAY>

A;Cross-references: UNIPROT:Q8X673; UNIPARC:UPI000000B85; GB:BA000007; PIDN:BAB35746.1;

A;Experimental source: strain O157:H7, substrain RIMD 0509952

C;Genetics:

A;Gene: EC52323

C;Superfamily: melibiose carrier protein

Query Match 6.8%; Score 81.5; DB 2; Length 457;

Best Local Similarity 25.2%; Pred. No. 11;

Matches 37; Conservative 30; Mismatches 53; Indels 27; Gaps 8;

QY 95 WDLRSILGAVFVSRTNNV-----VLEA-----PFLVGIEGSLKSTYNLL-FCGSCG 141
Db 317 WSLPVALVALATASIGQGVMTVMWALEADTVVEGYLTGVR--IEGLTYSLSFSTRKCG 374
QY 142 IPVGPHLYSTHAALAAALRGHFC--LSSDKMVCYLLKTKAIVNASEMDIQNV-----PLSE 194
Db 375 QAIG--GSIPAFILGLSGYIANQVTPVIMGIRTSIALVPCGFMLLAFVLIWFYPLTD 431
QY 195 KIAELKEKIVLTHNRKLSMKILSEVT 221
Db 432 K--KPKFIVEIDNRKKVQQQLISDIT 456

RESULT 15

B64918

glucuronide permease uidB - Escherichia coli (strain K-12)

C;Species: Escherichia coli

C;Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 09-Jul-2004

C;Accession: B64918

R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co

A.; Rose, D.J.; Mau, B.; Shao, Y.

Science 277, 1453-1462, 1997

A;Title: The complete genome sequence of Escherichia coli K-12.

A;Reference number: A64720; MUID:97426617; PMID:9278503

A;Accession: B64918

A;Status: nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-457 <BLAT>

A;Cross-references: UNIPROT:P30868; UNIPARC:UPI00000137B23; GB:AE000257; GB:U00096; NID:9

A;Experimental source: strain K-12, substrain MG1655

C;Genetics:

A;Gene: uidB; gusB

C;Superfamily: melibiose carrier protein

F;37-53/Domain: transmembrane #status predicted <TM01>

F;81-97/Domain: transmembrane #status predicted <TM02>

F;152-168/Domain: transmembrane #status predicted <TM03>

F;184-200/Domain: transmembrane #status predicted <TM04>

F;231-247/Domain: transmembrane #status predicted <TM05>

F;263-279/Domain: transmembrane #status predicted <TM06>

F;310-326/Domain: transmembrane #status predicted <TM07>

F;408-424/Domain: transmembrane #status predicted <TM08>

Query Match 6.8%; Score 81.5; DB 2; Length 457;

Best Local Similarity 25.2%; Pred. No. 11;

Matches 37; Conservative 30; Mismatches 53; Indels 27; Gaps 8;

QY 95 WDLRSILGAVFVSRTNNV-----VLEA-----PFLVGIEGSLKSTYNLL-FCGSCG 141
Db 317 WSLPVALVALATASIGQGVMTVMWALEADTVVEGYLTGVR--IEGLTYSLSFSTRKCG 374
QY 142 IPVGPHLYSTHAALAAALRGHFC--LSSDKMVCYLLKTKAIVNASEMDIQNV-----PLSE 194
Db 375 QAIG--GSIPAFILGLSGYIANQVTPVIMGIRTSIALVPCGFMLLAFVLIWFYPLTD 431
QY 195 KIAELKEKIVLTHNRKLSMKILSEVT 221
Db 432 K--KPKFIVEIDNRKKVQQQLISDIT 456

Search completed: March 25, 2006, 08:03:57.

Job time : 42 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 25, 2006, 07:56:15 ; Search time 231 Seconds
(without alignments)
699.420 Million cell updates/sec

Title: US-09-942-052A-728

Perfect score: 1198

Sequence: 1 MAAPQLRHRSRCATPPRGDF.....LKSLMKILSEVTPDQSKPEN 229

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

UniProt 05.80.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1198	100.0	229	1 OIP5 HUMAN	O43482 homo sapien
2	122	10.2	218	2 O9CXK6 MOUSE	O9CXK6 mus musculus
3	118	9.8	462	2 O4PAB5 USTMA	O4PAB5 ustilago ma
4	115.5	9.6	204	2 O9CZJ6 MOUSE	O9CZJ6 mus musculus
5	113	9.4	233	1 CU045 HUMAN	O9NYP9 homo sapien
6	113	9.4	233	2 O54220 HUMAN	O54220 homo sapien
7	110.5	9.2	232	1 CU045 PANTR	O68ut5 pan troglod
8	109	9.1	207	2 O5BLB1 BRARE	O5BLB1 brachydanio
9	107	8.9	532	2 O54HL6 DICDI	O54HL6 dictyoscell
10	105	8.8	155	1 YCZC SCHPO	O9P802 schizosacch
11	94	7.8	1556	2 O4Q626 LEIMA	O4Q626 leishmania
12	92	7.7	361	2 O8S611 ORYGA	O8S611 oryza sativ
13	92	7.7	695	2 O9CH87 LACLA	O9CH87 lactococcus
14	91	7.6	628	2 O69KE0 ORYSA	O69KE0 oryza sativ
15	90.5	7.6	597	2 O83693 TREPA	O83693 treponema p
16	89.5	7.5	502	2 O50P02 ENTPI	O50P02 entamoeba h
17	89.5	7.5	661	2 O50V43 ENTPI	O50V43 entamoeba h
18	89.5	7.5	661	2 O50W28 ENTPI	O50W28 entamoeba h
19	89	7.4	633	2 O5M052 STRT1	O5M052 streptococc
20	89	7.4	664	1 CHER HUMAN	O96ep1 homo sapien
21	89	7.4	735	2 O5M4E3 STRY2	O5M4E3 streptococc
22	89	7.4	893	1 POL2 BANWA	O65329 barley mild
23	88.5	7.4	152	2 O50XW7 ENTPI	O50XW7 entamoeba h
24	88.5	7.4	638	2 O75RZ6 CIOIN	O75RZ6 ciona intes
25	88	7.3	250	1 KKA7 CAME	P14508 campylobact
26	88	7.3	493	2 O68VH2 9POTY	O68VH2 barley mild
27	87.5	7.3	634	2 O6CX82 KLUKA	O6CX82 kluyveromyc
28	87.5	7.3	636	2 O4UTS4 CORUK	O4UTS4 corynebacte
29	87	7.3	664	1 CHER MOUSE	O81013 mus musculus
30	86.5	7.2	669	2 O97FQ5 CLOAB	O97FQ5 clostridium
31	86.5	7.2	2093	2 O5K2K3_9RHAB	O5K2K3 isfahan vir

32	86.5	7.2	5344	2 Q529N3 MAGGR	Q529N3 magnaporthe
33	86	7.2	145	2 Q839T4 ENTFA	Q839T4 enterococcu
34	86	7.2	327	2 Q4NE64 9MICC	Q4NE64 arthrobacte
35	86	7.2	331	2 O6GXE5 COCCA	O6GXE5 cochllobolu
36	86	7.2	893	2 O68VH4 9POTY	O68VH4 barley mild
37	86	7.2	3692	2 O5J1O7_9NOCA	O5J1O7 nocardia un
38	85	7.1	337	1 SW1 STRAW	Q82hul streptomyc
39	85	7.1	671	2 O815K5 PLAF7	O815K5 plasmodium
40	85	7.1	1018	2 O6FKN6 CANGA	O6FKN6 candida gla
41	85	7.1	2194	1 GLSN MEDSA	Q03460 medicago sa
42	85	7.1	2194	2 O40360 MEDSA	Q40360 medicago sa
43	84.5	7.1	361	2 O8YV34 ANASP	O8YV34 anabaena sp
44	84.5	7.1	376	2 O5YZE1 NOCFA	O5YZE1 nocardia fa
45	84	7.0	247	2 Q64WS0_BACFR	Q64WS0 bacteroides

ALIGNMENTS

RESULT 1

OIP5_HUMAN
ID OIP5_HUMAN STANDARD; PRT; 229 AA.
AC O43482; Q96BX7;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Opa-interacting protein 5.
GN Name=OIP5;
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=98125741; PubMed=9466265;
RA Williams J.M., Chen G.-C., Zhu L., Rest R.F.;
RT "Using the yeast two-hybrid system to identify human epithelial cell
proteins that bind gonococcal Opa proteins: intracellular gonococci
bind pyruvate kinase via their Opa proteins and require host pyruvate
for growth.";
RL Mol. Microbiol. 27:171-186(1998).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC TISSUE=Uterus;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,
Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
Rosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- SUBUNIT: Binds outer membrane protein Opa from Neisseria
gonorrhoeae.
CC -!- INTERACTION:
CC P04049:RAFI; NBExp=3; IntAct=EBI-536879, EBI-365996;
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.

DR EMBL: AF025441; AAC39561.1; ALT INIT; mRNA.
 DR EMBL: BC015050; AAHL5050.1; mRNA.
 DR InAct; O43482; ...
 DR Ensembl; ENSG00000104147; Homo sapiens.
 DR HGNC; HGNC:20300; OIP5.
 DR H-invDB; HIX0012150; ...
 DR MIA; 606020; ...
 DR GO; GO:0005515; F:protein binding; TAS.
 DR GO; GO:0007154; P:cell communication; NAS.
 DR GO; GO:0007154; P:cell communication; NAS.
 SQ SEQUENCE 229 AA; 24691 MW; 0EBD4006193A3106 CRC64;

Query Match 100.0%; Score 1198; DB 1; Length 229;
 Best Local Similarity 100.0%; Pred. No. 8.e-105;
 Matches 229; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MAAGPLHRRCATPPRCGCGTTERAIDQASFTTSMENDTQVVGSSPLGAGABEP 60
 DB 1 MAAGPLHRRCATPPRCGCGTTERAIDQASFTTSMENDTQVVGSSPLGAGABEP 60
 QY 61 AAGPQLPSWLOPERCAVFCQAQCHAVLADSVHLAWLDSRLGAVFGRVNNVLEAPFL 120
 DB 61 AAGPQLPSWLOPERCAVFCQAQCHAVLADSVHLAWLDSRLGAVFGRVNNVLEAPFL 120
 QY 121 VGIEGSLKSTYNLLFCGSCGIPVGFHLYSTHAALALRGHFLCLSSDKWCVYLLKTKAIV 180
 DB 121 VGIEGSLKSTYNLLFCGSCGIPVGFHLYSTHAALALRGHFLCLSSDKWCVYLLKTKAIV 180
 QY 181 NASEMDIQNVPLSEKIAELKEKIVLTHNRKLSMKILSEVTPDQSKPEN 229
 DB 181 NASEMDIQNVPLSEKIAELKEKIVLTHNRKLSMKILSEVTPDQSKPEN 229

RESULT 2

QY Q9CXK6_MOUSE PRELIMINARY; PRT; 218 AA.
 AC Q9CXK6; ...
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE Mus musculus 13 days embryo head cDNA, RIKEN full-length enriched
 DE library, clone:3110025H23 product:hypothetical protein, full insert
 DE sequence.
 GN Name=2610039C10Rik;
 OS Mus musculus (Mouse)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 OC Muridea; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RN NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Head;
 RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
 RA Carninci P.; Hayashizaki Y.;
 RT "High-efficiency full-length cDNA cloning";
 RL Meth. Enzymol. 303:19-44(1999).
 RN [2]
 RN NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Head;
 RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
 RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa K., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
 RA Saito T., Okazaki Y., Gotohori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gisi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schriml L.M., Staublei P., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,

RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-P.,
 RA Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohetsuki S.,
 RA Hayashizaki Y.;
 RA "Functional annotation of a full-length mouse cDNA collection";
 RL Nature 409:685-690(2001).
 RN [3]
 RN NUCLEOTIDE SEQUENCE
 RC STRAIN=C57BL/6J; TISSUE=Head;
 RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
 RA Carninci P., Shibata Y., Hayatsu M., Sugahara Y., Shibata K., Itoh M.,
 RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new genes";
 RL Genome Res. 10:1617-1630(2000).
 RN [5]
 RN NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Head;
 RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
 RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
 RA Konno H., Akiyama J., Nishi K., Kiteunai T., Tashiro H., Itoh M.,
 RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
 RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
 RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
 RT "RIKEN integrated sequence analysis (RISA) system-384-format
 sequencing pipeline with 384 multicapillary sequencer";
 RL Genome Res. 10:1757-1771(2000).
 RN [6]
 RN NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Head;
 RA Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,
 RA Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,
 RA Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,
 RA Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,
 RA Kawai J., Kojima Y., Konno H., Kouda M., Koya S., Kurihara C.,
 RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,
 RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,
 RA Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,
 RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,
 RA Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,
 RA Muramatsu M., Hayashizaki Y.;
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AK014084; BAB29147.1; -; mRNA.
 DR Ensembl; ENSMUSG0000022978; Mus musculus.
 DR MGI; MGI:1913828; 2610039C10Rik.
 KW Hypothetical protein.
 SQ SEQUENCE 218 AA; 24447 MW; 5539085BFDDDE8C2 CRC64;

Query Match 10.2%; Score 122; DB 2; Length 218;
 Best Local Similarity 23.3%; Pred. No. 0.0049;
 Matches 48; Conservative 30; Mismatches 102; Indels 26; Gaps 6;
 QY 29 DQASFTTSMENDTQVVGSSPLGAGABEPAGPQLPSWLOPERCAVFCQAQCHAVLA 88
 DB 15 DSSRYLRLOKWAN--MSSDALGLEKPERPEKAAAENP-----LVFLCARCRPLG 64
 QY 89 DSVHLAWLDSRLGAVFGRVNNVLEAPFLVGIEGSLKSTYNLLFCGSCGIPVGFH 147
 DB 65 DS--LTWASQEDTNCILLRSVCSNVSDKEPKLSKCRDEGDCILEALYCTGCSLSGVY 122

QY 148 LYSTHAALAAALRGHFCISDRKVCYLL-----KTKAIIVNASEMDIQNVPLSEKTAIE 198
 DB 123 YRCTPKNLDYKRDLCFLSVETVYSLGSSEKQIVSKDKELFNLE-----SRVEIEKSIQ 178
 QY 199 LKEKIVLTHNRLKSLMKLILSEVTPDQ 224
 DB 179 MEEVITALQKLRVESKLSLAQPCQ 204

RESULT 3

Q4PAB5 USTWA
 ID Q4PAB5 USTWA PRELIMINARY; PRT; 462 AA.
 AC Q4PAB5
 DT 13-SEP-2005 (TrEMBLrel. 31, Created)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
 DE Hypothetical protein.
 GN ORFNames=UM02948.1;
 OS Ustilago maydis 521.
 OC Eukaryota; Fungi; Basidiomycota; Ustilaginomycetes;
 OC Ustilaginomycetidae; Ustilaginales; Ustilaginaceae; Ustilago.
 OX NCBI_TaxID=237631;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.

RC STRAIN=521;
 RA Birren B., Nusbaum C., Abebe A., Abouelleil A., Adekoya E.,
 RA Alt-zahra M., Allen N., Allen T., An P., Anderson M., Anderson S.,
 RA Arachi H., Armbruster J., Bachtasang P., Baldwin J., Barry A.,
 RA Bayul T., Blitshsteyn B., Bloom T., Blye J., Boguslavskiy L.,
 RA Borowsky M., Boukhgaltier B., Brunache A., Butler J., Calixte N.,
 RA Calvo S., Camarata J., Campo K., Chang J., Cheshatsang Y., Citroen M.,
 RA Callimore A., Considine T., Cook A., Cooke P., Corum B., Cuomo C.,
 RA David K., Dawoe T., Degray S., Dodge S., Dooley K., Dorje P.,
 RA Dorjee K., Dorris L., Duffey N., Dupes A., Elkins T., Engels R.,
 RA Erickson J., Farina A., Faro S., Ferreira P., Fischer H.,
 RA Fitzgerald M., Foley K., Gage D., Galagan J., Gearin G., Gerner S.,
 RA Girke A., Goyette A., Graham J., Grandbois E., Gyaltsen K., Hafez N.,
 RA Hagopian D., Hagos B., Hall J., Hatcher B., Heiler A., Higgins H.,
 RA Honan T., Horn A., Houde N., Hughes L., Hulme W., Hueby E., Iliev I.,
 RA Jaffe D., Jones C., Kamal M., Kamat A., Kamvesselis M., Karlsson E.,
 RA Kells C., Kieu A., Kisser P., Kodira C., Kulbokas E., Labutti K.,
 RA Lama D., Landers T., Leger J., Levine S., Lewis D., Lewis T.,
 RA Lindblad-toh K., Liu X., Lokyitsang T., Lokyitsang Y., Lucien O.,
 RA Lui A., Ma L.J., Mabbitt R., Maru K., Matthews J., Maclean C., Major J.,
 RA Manning J., Maraballa R., Mcghee T., Meldrim J., Meneus L.,
 RA McCarthy M., McDonough S., Mcghee T., Meldrim J., Meneus L.,
 RA Mesirov J., Mihalev A., Mihova T., Mikkelsen T., Mlenga V., Moru K.,
 RA Mozes J., Mulrain L., Munson G., Naylor J., News C., Nguyen C.,
 RA Nguyen N., Nguyen T., Nicol R., Nielsen C., Nizzari M., Norbu C.,
 RA Norbu N., O'donnell P., Okoawo O., O'leary S., Omotoho B.,
 RA O'neil K., Osman S., Parker S., Perrin D., Phunkhang P., Piquani B.,
 RA Purcell S., Rachupka T., Ramasamy U., Rameau R., Ray V., Raymond C.,
 RA Retta R., Richardson S., Rise C., Rodriguez J., Rogers J., Rogov P.,
 RA Rutman M., Schupbach R., Seaman C., Settipalli S., Sharpe T.,
 RA Sheridan J., Sherpa N., Shi J., Smirnov S., Smith C., Sougnev C.,
 RA Spencer B., Stalker J., Stange-thomann N., Stavropoulos S.,
 RA Stetson K., Stone C., Stone S., Stubbs M., Talamas J., Tchuinga P.,
 RA Tenzing P., Tefaye S., Theodore J., Thoultsang Y., Topham K.,
 RA Towey S., Teasla T., Tsomo N., Vallee D., Vassiliev H.,
 RA Venkataraman V., Vinson J., Vo A., Wade C., Wang S., Wangchuk T.,
 RA Wangdi T., Whitaker C., Wilkinson J., Wu Y., Wyman D., Yadav S.,
 RA Yang S., Yang X., Yeager S., Yee E., Young G., Zainoun J., Zembeck L.,
 RA Zimmer A., Zody M., Zander E.;

RT "The genome sequence of Ustilago maydis.";
 RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
 CC -!- CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 DR ENBL; AACP01000101; EAK84120.1; -; Genomic DNA.
 KW Hypothetical protein.
 SQ SEQUENCE 462 AA; 49283 MW; EDDCFDF17757B509 CRC64;

Query Match 9.8%; Score 118; DB 2; Length 462;
 Beat Local Similarity 25.9%; Pred. NO. 0.03;
 Matches 64; Conservative 28; Mismatches 109; Indels 46; Gaps 10;
 QY 2 AAQPLRHRSRCATPRGDFCGTERTAIQAASFTTSMENDTQV-VKGSSPLGAGLGAEP 60
 DB 109 AAQPLSSDDDDMAPPKG-----GSLTTT---STRARGRGGRGPRGRGRG 153
 QY 61 AA-----GPOL--PSWLOPERCAVFCQAACHAVLADSV-HLAWDLSSRLGAVVFSRYT 110
 DB 154 ASTSTTVVHKPDLESDESEPPNPLVFCQRCCLGSLAFVATDI--DLGVILSDVS 211
 QY 111 NNVLREAPPLVGIE-GSLKSGSTYNLLFCGSGCIPVGFHYLSTHAALALRGHFCISSDQM 169
 DB 212 EIIQODTTYETSTETPGKDIQTFARLCAGCAAAVGRNYRTTPRDLDLRDCFSLEVDI 271
 QY 170 VCYLLKTKAIIVNASEMDIQ-----NVPLSEKIAELKEKIVLTH-NRLKS 212
 DB 272 YTYQLSGSNYTRQKEEDDEDQDALVDGGASGKPARASRPTHETVQAEDTKVLITKMER 331
 QY 213 LMKILSE 219
 DB 332 LTIELSD 338

RESULT 4

Q9CZJ6 MOUSE
 ID Q9CZJ6 MOUSE PRELIMINARY; PRT; 204 AA.
 AC Q9CZJ6;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
 DE Mus musculus 11 days embryo whole body cDNA, RIKEN full-length
 DE enriched library, clone:2700078124 product:hypothetical protein, full
 DE insert sequence (2610039C10Rik protein).
 GN Name=2610039C10Rik;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 OC Muroidae; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP NUCLEOTIDE SEQUENCE
 RC STRAIN=C57BL/6J; TISSUE=Whole body;
 RC MEDLINE=92279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
 RA Carninci P., Hayashizaki Y.;
 RT "High-efficiency full-length cDNA cloning.";
 RN Meth. Enzymol. 303:19-44(1999).
 RN [2]

RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Whole body;
 RC MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
 RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Azakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana I.,
 RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,
 RA Kadoya K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schram L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barch G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.P.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seva T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kotsuki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RN Nature 409:685-690(2001).
 RN [3]

RA	Taenzler S., Galgoczy P., Platzer M., Scharfe M., Nordsieck G.,
RA	Bloeker H.J., Hellmann I., Khaitovich P., Paabo S., Reinhardt R.,
RA	Zheng H.-J., Zhang X.-L., Zhu G.-F., Wang B.-F., Fu G., Ren S.-X.,
RA	Zhao G.-P., Chen Z., Lee Y.-S., Cheong J.-E., Choi S.-H., Wu K.-M.,
RA	Liu T.-T., Heiao K.-J., Teai S.-P., Kim C.-G., Ooi S.-K., Kitano T.,
RA	Kohara Y., Saitou N., Park H.-S., Wang S.-Y., Yaspo M.-L., Sakaki Y.,
RT	"DNA sequence and comparative analysis of chimpanzee chromosome 22,"
RL	Nature 429:382-388(2004).
CC	-/- SIMILARITY: To S.pombe SpC970.12.
CC	-----
CC	This Swiss-Prot entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use as long as its content is in no way modified and this statement is not
CC	removed.
CC	-----
DR	EMBL; ALJ954205; CAHI8576.1; -; Genomic DNA.
SQ	SEQUENCE 232 AA; 25832 MW; 6549B8BE36924587 CRC64;

	Query Match 9.2%; Score 110.5; DB 1; Length 232;
	Best Local Similarity 27.1%; Pred. No. 0.065;
	Matches 52; Conservative 29; Mismatches 88; Indels 23; Gaps 8;
OY	31 ASPTTSMEWDTQVVGSSPLGPAGLAGAEPAAGQLPSWLQPRCAVFQCAQCHAVLAOS 90
DB	49 ASMWSMSDASV-----ADMERARL--EEAAAA-----EERPLVFLCSGCCRRPLGDS 94
OY	91 VHLAWDLRS-SLGNAVFSRVTNNVLLEAPLVGLIEGSLKGSTYNLLFCGSCGIPVGFHLY 149
DB	95 --LSWVASQEDTNCILRCVSNVDVKQKLKRKENGCVLETCCAGCSLNULGVTR 152
OY	150 STHAALAAARGHFCLSSDDNVCYLL-KTKAIYNASEMDIQNVPLSEKIAELKEKIVLTH 207
DB	153 CTPEKNLDYKRDLCFLSVESATSYVIGSGSEQIV-SEDKELFNL--ESRVEIEKSLTOME 208
OY	208 NRLLSKMLKIUSE 219
DB	209 DVLKALQMKLWE 220
RESULT 8	
QSBLBI_BRARE	
ID	QSBLBI_BRARE PRELIMINARY; PRT; 207 AA.
AC	QSBLBI;
DC	10-MAY-2005 (TEMBLrel. 30, Created)
DT	10-MAY-2005 (TEMBLrel. 30, Last sequence update)
DT	10-MAY-2005 (TEMBLrel. 30, Last annotation update)
DE	LOC553502 protein (Fragment).
GN	Name=LOC553502;
OS	Brachydanio rerio (Zebrafish) (Danio rerio).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OX	Cyprinidae; Danio.
NCBI_TaxID=7955;	
RN	[1]
NCLEOTIDE SEQUENCE.	
RP	STRAIN=Singapore local strain; TISSUE=Embryo;
RX	MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.F., Schuler G.D.,
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.M., Bhat N.K.,
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA	Scapleton M., Soares J.B., Bonaldo M.F., Casavant T.B., Scheetz T.E.,
RA	Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mallahy S.J.,
RA	Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA	Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.
RA	Fahay J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA	Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,

```
SQ SEQUENCE 532 AA; 60715 MW; BAF75AB8CD7C1D5C CRC64;
Query Match 8.9%; Score 107; DB 2; Length 532;
Best Local Similarity 22.9%; Pred. No. 0.4;
Matches 38; Conservative 29; Mismatches 75; Indels 24; Gaps 5;
QY 77 VFQCAQCHAVLADSVHLAWDLRSGLGAVFVSRTNNVLEAPLVGIEGSLKST-----131
D 205 IFSCLSCKTPIGDSTLI---VNSKHECDMMYLTNTFVWKLGSGVILGDQKPESTYRGST 261
QY 132 ---YNLLFCGSCGIPVGFHLYSTHAALALRGHFCLSDDKMCVYLL-KTKAIVNASEMDI 187
D 262 DCIYSILFYSCYNPLGRKYNKTSOKFKILNNHFVVDIGSITYSYSGNNNIIIVDKNEI 321
QY 188 QNVPLS-----EKIABLKEKIVLTHRLKSLMKILSEVTPD 223
D 322 ---VPLSFGNDKLTQMHTNMWNEKINFLEDOQLLSVFNISIKALPE 365
RESULT 10
YCZC SCHPO
ID YCZC SCHPO STANDARD; PRT; 155 AA.
AC Q9P802;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Hypothetical protein C970.12 in chromosome III.
GN ORFNames=SPCC970.12;
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX MEDLINE=21848401; PubMed=11859360; DOI=10.1038/nature724;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M.H., Lyne R., Stewart A.,
RA Sgouros J.G., Peat N., Hayles J., Baker S.G., Basham D., Bowman S.,
RA Brooks K., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D.E., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K.D., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K.L., Murphy L.D., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford J.K.M., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M.N., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J.R., Volckaert G., Aert R., Robben J., Grynoprez B.,
RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Duesterhoeft A., Fritze C., Holzer E., Moestl D.,
RA Hilbert H., Borzym K., Langer I., Beck A., Leirach H., Reinhardt R.,
RA Pohl T.M., Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert P., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Sipakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe.";
RL Nature 415:871-880(2002).
CC -!- SIMILARITY: To human C21orf45.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
CC EMBL; AL031530; CAB72327.1; -; Genomic_DNA.
CC GenedB_Spombe; SPCC970.12; -;
```

```
DR GO; GO:0005515; F:protein binding; IPI.
DR GO; GO:0031055; P:chromatin remodeling at centromere; IMP.
DR GO; GO:0007059; P:chromosome segregation; IMP.
DR GO; GO:0016575; P:histone deacetylation; IMP.
DR GO; GO:0031066; P:regulation of histone deacetylation at cent.; IMP.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 155 AA; 17874 MW; A7ABBD0F93760B0B CRC64;
Query Match 8.8%; Score 105; DB 1; Length 155;
Best Local Similarity 30.1%; Pred. No. 0.13;
Matches 41; Conservative 18; Mismatches 61; Indels 16; Gaps 6;
QY 71 QPERCAVFOCAQCHAVLADSVHLAWDLRS--LGAVFVSRTNNVLEAPLVGIEGSLK 128
D 19 QP---SVFQCKKCFQIVGDS--NAWVISHREYLSFTLSDAVENSVRVEDTFKSDGGL-- 71
QY 129 GSTYNLLFCGSCGIPVGFHLYSTHAALALRGHFCLSDDKMCVYLLKTKAIVNASEMDI 188
D 72 -CVSELSCTRCNEVIGKYNSTPYLDDIRDMYTFMDKQLQAYQLGNKT-VNPEGLTRY 129
QY 189 NVPLSEKIAELKEKIV 204
D 130 QVDL-----EMREDII 140
RESULT 11
Q4Q626 LEIMA
ID Q4Q626 LEIMA PRELIMINARY; PRT; 1556 AA.
AC Q4Q626;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=LmjF31.2350;
OS Leishmania major.
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5664;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Peacock C.S., Murphy L., Ivens A.C., Berriman M., Blackwell J.,
RA Smith D., Collins M., Posker N., Harris D., Oliver K., O'Neill S.,
RA Saunders D., Seeger K., Warren T., Rajandream M., and Barrell B.G.;
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; CT005268; CAJ08424.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 1556 AA; 170195 MW; F6FF7CDDF8DCA59 CRC64;
Query Match 7.8%; Score 94; DB 2; Length 1556;
Best Local Similarity 24.3%; Pred. No. 27;
Matches 45; Conservative 22; Mismatches 50; Indels 68; Gaps 9;
QY 22 GGTERAIDQASFTTSMSE-----WDTQVVKGS-----SPLGPAGLGAEEPAAGPQ 65
D 833 GGGELMQTDAGFSEVESPIAFLYDDDDVHRASQDVQKREEPDGVGGAGEDAAAG--- 890
QY 66 LPSWLQPERCAVFOCAQCHAVLADSV-----HLAWDLRSGLGAVFVSRTNNVLEAP 118
D 891 -----ARAAGE--QDFGLNDTVAAPSTYSHMLCRVLELIRLTARLNILLDLVLOAP 941
QY 119 FLVIGIEGSLKSTYNLLFCGSCGIPVGFHLYSTHAALALRGHFCLSDDKMCVYLL 174
D 942 VL-----TRY-----VWYHLK-----HFCLSDAKASCVIGMWLL 971
QY 175 KTKAI 179
D 972 RNLAIV 976
RESULT 12
Q8S611 ORYSA
ID Q8S611 ORYSA PRELIMINARY; PRT; 361 AA.
AC Q8S611;
```


DT 01-JUN-2002 (TREMREL. 21, Created)
DT 01-JUN-2002 (TREMREL. 21, Last sequence update)
DT 01-FEB-2005 (TREMREL. 29, Last annotation update)
DE Putative gypsy-type retrotransposon protein.
GN ORFNames=OSJNB0096R22.5;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Eriarthoideae; Oryzae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Wang R.A., Yu Y., Soderlund C., Chen M., Kim H.-R., Rambo T.,
RA Sasaki C., Henry D., Oates R., Simmons J.,
RA Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
RL [2]
RP NUCLEOTIDE SEQUENCE.
RA The Rice Chromosome 10 Sequencing Consortium;
RT "In-depth view of structure, activity, and evolution of rice
RT chromosome 10.";
RL Science 300:1566-1569 (2003).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RA Buell C.R., Wing R.A., McCombie W.R., Messing J., Yuan Q.,
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
RL EMBL; AC094400; AAL91599.1; -; Genomic_DNA.
DR EMBL; AE017047; AAP51763.1; -; Genomic_DNA.
DR Gramene; Q8S611; -;
SQ SEQUENCE 361 AA; 38804 MW; F883B3E8B8FF45 CRC64;

Query Match 7.7%; Score 92; DB 2; Length 361;
Best Local Similarity 25.3%; Pred. No. 6.4;
Matches 61; Conservative 28; Mismatches 92; Indels 60; Gaps 14;
Qy 2 AAQPLHRSRCATPPRGDFCGGTGTERAIDQASFTTSMEDWTQVKGSSPLGAGLCABEPA 61
Db 49 APTLPFRRAVRAKAAQDGGTSSPAVAST-----DVVVFGSREATPSG-PASDPV 102
Qy 62 AGPOLP-----SW--LQPRCAVFOCAQCHAVLADSLRSLSGLGAVFVRVNNVL 115
Db 103 AGRSPAAVLWSBELQVEMGRLEAGA--RVIGREIAEARGLEHRM-----SELGN-- 152
Qy 116 EAPFLVIGSLGKSTYNLL--FCGSGI-----PVGFHYLTHAALAA-----RGH 161
Db 153 ----LSEIRGLR-VTYTGLHOLAGKCGIKTIPANPDFEFLTSLAEALAAEMEISKH 207
Qy 162 FCLSSDKM-----VCYLLKTKAIVNASMDIQNV-----PLSEKIAELKEKI 203
Db 208 AARIGEEMSNRIYIGACHILACVRLAH-PELDRLREILDQGEASDARKDVMEEVGLGKSV 266
Qy 204 V 204
Db 267 L 267

RESULT 13
Q9CH87_LACLA
ID Q9CH87_LACLA PRELIMINARY; PRT; 695 AA.
AC Q9CH87;
DT 01-JUN-2001 (TREMREL. 17, Created)
DT 01-JUN-2001 (TREMREL. 17, Last sequence update)
DT 01-MAR-2004 (TREMREL. 26, Last annotation update)
DE Copper-potassium transporting ATPase B.
GN Name=copB; OrderedLocNames=LL0851;
OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.
OX NCBI_TaxID=1360;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA STRAIN=ILI403;
RX MEDLINE=21235186; PubMed=11337471; DOI=10.1101/gr-1697R;
RA Bolotin A., Wincker P., Manger S., Jaillon O., Malarne K.,
RA Weissenbach J., Ehrlich S.D., Sorokin A.;

RT "The complete genome sequence of the lactic acid bacterium Lactococcus
RT lactis ssp. lactis IL1403.";
RL Genome Res. 11:731-753 (2001).
DR EMBL; AE006319; AAK04949.1; -; Genomic_DNA.
DR PIR; C86731; C86731.
DR GO; GO:0016021; C: integral to membrane; IEA.
DR GO; GO:0016020; C: membrane; IEA.
DR GO; GO:0005524; F: ATP binding; IEA.
DR GO; GO:0003824; F: catalytic activity; IEA.
DR GO; GO:0004008; F: copper-exporting ATPase activity; IEA.
DR GO; GO:0016820; F: hydrolase activity, acting on acid anhydrid. .; IEA.
DR GO; GO:0046872; F: metal ion binding; IEA.
DR GO; GO:0046873; F: metal ion transporter activity; IEA.
DR GO; GO:0008152; P: metabolism; IEA.
DR GO; GO:0030001; P: metal ion transport; IEA.
DR GO; GO:0015992; P: proton transport; IEA.
DR InterPro; IPR006403; ATPase-IB1_Cu.
DR InterPro; IPR006416; ATPase-IB_hvy.
DR InterPro; IPR001757; ATPase-IB_E2.
DR InterPro; IPR005834; Dehal_Like_Hydro.
DR InterPro; IPR008250; E1-E2_ATPase_reg.
DR Pfam; PF00122; E1-E2_ATPase; 1.
DR Pfam; PF00702; Hydrolase; 1.
DR PRINTS; PR00119; CATATPASE.
DR TIGRFAMs; TIGR01511; ATPase-IB1_Cu; 1.
DR TIGRFAMs; TIGR01525; ATPase-IB_hvy; 1.
DR TIGRFAMs; TIGR01494; ATPase-P-type; 2.
DR PROSITE; PS00154; ATPASE_E1_E2; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 695 AA; 75981 MW; 1B4947C3C2A0FA0F CRC64;

Query Match 7.7%; Score 92; DB 2; Length 695;
Best Local Similarity 25.3%; Pred. No. 15;
Matches 38; Conservative 28; Mismatches 50; Indels 34; Gaps 7;
Qy 101 LGAVFVRVNNVLAPFLVIGSLGKSTYNLLFCGSGIPIVGFHYLTHAALAAALRG 160
Db 101 LGTIIP-----FVSGTFFSGAGKELSKRPAMMLITWGITVAY-AYSVYAIMSLG 153
Qy 161 HF-----CLSSDKMVCYLLKTKAIVNASB-MDI-----QNVPLSR-K 195
Db 154 HMGNFWFELATLIVIMLIGHLEIKAMGAGDALKDASLVPPKKAHLKSGKDVLSBLK 213
Qy 196 IAELEKELVLTNRLKSLMKILSRVTPQS 225
Db 214 VGDL--LLVKNKIPADGLILSEALVDES 241

RESULT 14
Q69KE0_ORYSA
ID Q69KE0_ORYSA PRELIMINARY; PRT; 628 AA.
AC Q69KE0;
DT 25-OCT-2004 (TREMREL. 28, Created)
DT 25-OCT-2004 (TREMREL. 28, Last sequence update)
DT 25-OCT-2004 (TREMREL. 28, Last annotation update)
DE Myosin heavy chain-like.
GN Name=P0023E10.17;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Eriarthoideae; Oryzae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Sasaki T., Matsumoto T., Hattori M., Sakaki Y., Katayose Y.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 9, PAC
RT clone:P0023E10.";
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP005934; BAD36600.1; -; Genomic_DNA.
DR Gramene; Q69KE0; -;
DR InterPro; IPR005829; Sug transporter.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; UNKNOWN_1.
SQ SEQUENCE 628 AA; 68281 MW; 3230B108C6061027 CRC64;

Query Match 7.6%; Score 91; DB 2; Length 628;
Best Local Similarity 23.7%; Pred. No. 16;
Matches 66; Conservative 34; Mismatches 82; Indels 96; Gaps 17;
QY 8 HRSRCATPRG-----DPCGTERAIDQASFTTSMEDWTDVVKGSSPLGPA 57
DB 141 YRKAPTSPSRVARAKAQAQDSGTSASPAVAST-----DVVVFGSREAAFGSL-A 194
QY 58 EEPAGPOLP-----SW--LQPRCAVFOC-----AQHAVLADSVH-----LAWDL 97
DB 195 SDLVAGPGSPAALVTWEELOVEMGRLEAGARGIGREIAEAPAAAAAASANERADR 254
QY 98 S-----RSLGAV-----PSRVTVNVLEAPFLVIGIEGSLKSTYNLL--P 136
DB 255 ABAREDLKMRRELVAQNRQOGLHRMSELENN-----LSEIRDSLR-VITYTGLHQL 306
QY 137 CGSCGI-----PVGFHYLSTHAAAL-----RGHFCLSSDKM-----VCYLLKTK 177
DB 307 AGECKIKSTIPANPDEFSLTSSLAELAAAMEEIPSKHAARTRETSNGIYTGACHVLACL 366
QY 178 AIVNAEMDIQNV-----PLSEKIAELKEKIV 204
DB 367 RL VH-PELDRLREILDQGAASDARKQDVMEVGDGLESVL 403

RESULT 15

ID 083693 TREPA PRELIMINARY; PRT; 597 AA.
AC 083693;
DT 01-NOV-1998 (TremBLrel. 08, Created)
DT 01-NOV-1998 (TremBLrel. 08, Last sequence update)
DT 01-JUN-2003 (TremBLrel. 24, Last annotation update)
DE Phosphoribosylglycinamide formyltransferase, putative.
GN OrderedLocusNames=TP0695;
OS Treponema pallidum.
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Treponema.
OX NCBI_TaxID=160;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Nichols;
RX MEDLINE=9832770; PubMed=965876; DOI=10.1126/science.281.5375.375;
RA Fraser C.M., Norris S.J., Weinstock G.M., White O., Sutton G.G.,
RA Dodson R.J., Gwinn M.L., Hickey E.K., Clayton R.A., Ketchum K.A.,
RA Sodergren E., Hardham J.M., McLeod M.P., Salzberg S.L., Peterson J.D.,
RA Khalak H.G., Richardson D.B., Howell J.K., Chidambaram M.,
RA Utterback T.R., McDonald L.A., Artiach P., Bowman C., Cotton M.D.,
RA Fujii C., Garland S.A., Hatch B., Horst K., Roberts K.M., Sandusky M.,
RA Weidman J.F., Smith H.O., Venter J.C.;
RT "Complete genome sequence of Treponema pallidum, the syphilis
spirochete";
RL Science 281:375-388(1998).
DR EMBL; AE001243; AAC65662.1; -; Genomic_DNA.
DR PIR; D71293; D71293.
DR TIGR; TP0695;
DR GO; GO:0016874; F:ligase activity; IEA.
DR InterPro; IPR00135; ATP-grasp.
DR InterPro; IPR011761; ATP_GRASP.
DR Pfam; PF02222; ATP-grasp; 1.
DR DR PROSITE; PS00975; ATP_GRASP; 1.
KW Complete proteome.
SQ SEQUENCE 597 AA; 63326 MW; 7B775117C9BFD6A6 CRC64;

Query Match 7.6%; Score 90.5; DB 2; Length 597;
Best Local Similarity 28.0%; Pred. No. 17;
Matches 51; Conservative 19; Mismatches 47; Indels 65; Gaps 12;
QY 12 CATPPRGDFCGGTERAID-----QASFT-----TSNEWPTQVKGSSPLGPA 53
DB 90 CALP--GHRLEATKNATKTRMRCFTRLRRCPRFTFLFPDPSFANDT-----PPGHA 140
QY 54 GLGAEEPAAGPQLPSWLOP--ERCAVFOC--AQHAVLADSVHAWDLRSILGAVVFSRVT 110

DB 141 RLCSHLHSGAGLSFPLVVKPTDNMGARGCTLAQCKDTLINACAVARQFSRS----- 190
QY 111 NNVLLEAPFLVIGIEGSLKSTYNLLFCGSCGIPVGFHYLSTHAAALRGHFCLSSDKM 170
DB 191 GRVIEE-EFIVGREFSLSG-----LIFDGT-----LYVT--ALA-----DRHI 225
QY 171 CY 172
DB 226 CF 227

Search completed: March 25, 2006, 08:03:11
Job time : 234 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 25, 2006, 08:03:30 ; Search time 47 Seconds
(without alignments)
402.824 Million cell updates/sec

Title: US-09-942-052A-728
Perfect score: 1198
Sequence: 1 MAAPLHRSRCATPRGDP.....LKSLMKILSEVTPDQSKPEN 229

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Issued Patents AA:*
1: /cgn2_6/ptodata/1/aaa/5 COMB.pap.*
2: /cgn2_6/ptodata/1/aaa/6 COMB.pap.*
3: /cgn2_6/ptodata/1/aaa/7 COMB.pap.*
4: /cgn2_6/ptodata/1/aaa/8 COMB.pap.*
5: /cgn2_6/ptodata/1/aaa/9 COMB.pap.*
6: /cgn2_6/ptodata/1/aaa/10 COMB.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	86	7.2	158	2	US-09-134-000C-6596
2	83	6.9	457	6	Sequence 6596, Ap
3	81.5	6.8	456	6	Patent No. 5268463-7
4	81.5	6.8	457	1	Sequence 6, Appli
5	81.5	6.8	457	1	Sequence 6, Appli
6	80.5	6.7	1711	1	Sequence 2, Appli
7	78.5	6.6	325	1	Sequence 3, Appli
8	78.5	6.6	325	1	Sequence 12, Appli
9	78.5	6.6	325	1	Sequence 12, Appli
10	78.5	6.6	325	2	Sequence 12, Appli
11	78.5	6.6	452	2	Sequence 12, Appli
12	78	6.5	190	2	Sequence 12, Appli
13	77	6.4	607	2	Sequence 12, Appli
14	77	6.4	719	1	Sequence 1, Appli
15	77	6.4	719	1	Sequence 3, Appli
16	77	6.4	719	2	Sequence 3, Appli
17	77	6.4	738	6	Sequence 3, Appli
18	76.5	6.4	776	2	Sequence 3, Appli
19	76	6.3	351	2	Sequence 3, Appli
20	76	6.3	1001	2	Sequence 3, Appli
21	75.5	6.3	310	1	Sequence 3, Appli
22	75.5	6.3	310	1	Sequence 3, Appli
23	75.5	6.3	310	1	Sequence 3, Appli
24	75.5	6.3	310	2	Sequence 3, Appli
25	75.5	6.3	310	2	Sequence 3, Appli
26	75.5	6.3	1449	2	Sequence 3, Appli
27	75.5	6.3	1723	2	Sequence 3, Appli

28 75.5 6.3 4545 1 US-08-804-227C-14 Sequence 14, Appli
29 75.5 6.3 4550 1 US-08-804-227C-8 Sequence 8, Appli
30 75.5 6.3 4550 1 US-08-804-198-2 Sequence 2, Appli
31 75 6.3 78 2 US-09-270-767-43496 Sequence 43496, A
32 75 6.3 458 2 US-09-800-170-3 Patent No. 5262177
33 75 6.3 717 6 5262177-5
34 75 6.3 784 2 US-09-724-586A-1 Sequence 1, Appli
35 75 6.3 784 2 US-09-600-823-1 Sequence 1, Appli
36 75 6.3 784 2 US-09-724-666-1 Sequence 1, Appli
37 75 6.3 1088 2 US-09-920-804-2 Sequence 2, Appli
38 75 6.3 1088 2 US-09-920-804-4 Sequence 4, Appli
39 75 6.3 1129 2 US-09-734-674-2 Sequence 2, Appli
40 75 6.3 1129 2 US-10-274-990-2 Sequence 2, Appli
41 75 6.3 2227 2 US-08-475-886-4 Sequence 4, Appli
42 75 6.3 2227 2 US-09-653-499-4 Sequence 4, Appli
43 75 6.3 2227 2 US-10-135-988-4 Sequence 4, Appli
44 74.5 6.2 211 2 US-09-949-016-11409 Sequence 11409, A
45 74.5 6.2 279 2 US-09-107-433-3335 Sequence 3335, Ap

ALIGNMENTS

RESULT 1
US-09-134-000C-6596
; Sequence 6596, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; PRIOR FILING DATE: 1998-08-13
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 6596
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-134-000C-6596

Query Match 7.2%; Score 86; DB 2; Length 158;
Best Local Similarity 26.5%; Pred. No. 0.18;
Matches 40; Conservative 27; Mismatches 54; Indels 30; Gaps 9;
Qy 87 LADSVHLANDLSRLGAVV---FSRVNNVVLEA--PFLVGIEGSLKG-----STVN 133
Db 17 LTEYIYLASEQPLSTGSIITLGLSKPTGNLKAFFFFINDEGDPGKFSYNYLKTILN 76
Qy 134 LLFCGSGGIPGVPHLYSTHAALAL---RGHFCLSSDRWCYLLKTKAIVNASEMDIONV 190
Db 77 YQVCSMCSLSDNQNIETPSAALLGVIRNHF-----KNTC---TTKUYILSALNGFENK 128
Qy 191 PLSE-----KIAELKEKIVLTHNRKLSMKIL 217
Db 129 PISKTVTRKLSLSYK-DLYVQEMK-LLKII 157

RESULT 2
5268463-7
; Patent No. 5268463
; APPLICANT: JEFFERSON, RICHARD A.
; TITLE OF INVENTION: PLAANT PROMOTER a-GLUCURONIDASE GENE
; CONSTRUCT
; NUMBER OF SEQUENCES: 9
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/447,976
; FILING DATE: 08-DEC-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 119,102

; FILING DATE: 10-NOV-1987
; APPLICATION NUMBER: 264,586
; FILING DATE: 31-OCT-1988
; SEQ ID NO: 7;
; LENGTH: 457
5268463-7

Query Match 6.9%; Score 83; DB 6; Length 457;
Best Local Similarity 25.2%; Pred. No. 2;
Matches 37; Conservative 30; Mismatches 54; Indels 26; Gaps 8;

QY 95 WDLRSGLGAVFSRVNNV-----VLEA-----PFLVGIEGSLKSTYNLL-FCGSCG 141
Db 316 WSLPVALVALAIAASIGQGVTTVMWALEADTVEYGYLTGVR--IEGLTYSLSFSTRKCG 373

QY 142 IPVGPHLYSTHAALALRGHFC--LSSDKMVCYLLKTKAIVNASMDIQNV-----PLSE 194
Db 374 QAIGGS--SIPAFILGSLGYIANQVTPVINGIRTSIALVPCGFMLLAFVIWFPYPLTD 431

QY 195 KIAELKEKIVLTHNRKLSMKILSEVT 221
Db 432 K--KFEIVVEIDNRKKVQQQLISDIT 456

RESULT 3

5432081-7
; Patent No. 5432081
; APPLICANT: JEFFERSON, RICHARD A.
; TITLE OF INVENTION: HOST CELLS TRANSFORMED WITH THE E. COLI
; GLUCORONIDE PERMEASE GENE
; NUMBER OF SEQUENCES: 10
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/138,546
; FILING DATE: 15-OCT-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 447,976
; FILING DATE: 08-DEC-1989
; APPLICATION NUMBER: 264,586
; FILING DATE: 31-OCT-1988
; APPLICATION NUMBER: 119,102
; FILING DATE: 10-NOV-1987
; SEQ ID NO: 7;
; LENGTH: 456
5432081-7

Query Match 6.8%; Score 81.5; DB 6; Length 456;
Best Local Similarity 25.2%; Pred. No. 2.9;
Matches 37; Conservative 30; Mismatches 53; Indels 27; Gaps 8;

QY 95 WDLRSGLGAVFSRVNNV-----VLEA-----PFLVGIEGSLKSTYNLL-FCGSCG 141
Db 316 WSLPVALVALAIAASIGQGVTTVMWALEADTVEYGYLTGVR--IEGLTYSLSFSTRKCG 373

QY 142 IPVGPHLYSTHAALALRGHFC--LSSDKMVCYLLKTKAIVNASMDIQNV-----PLSE 194
Db 374 QAIG--GSIPAFILGSLGYIANQVTPVINGIRTSIALVPCGFMLLAFVIWFPYPLTD 430

QY 195 KIAELKEKIVLTHNRKLSMKILSEVT 221
Db 431 K--KFEIVVEIDNRKKVQQQLISDIT 455

RESULT 4

US-08-882-704A-6
; Sequence 6, Application US/08882704A
; Patent No. 5879906
; GENERAL INFORMATION:
; APPLICANT: Jefferson, Richard A.
; APPLICANT: Wilson, Katherine J.
; TITLE OF INVENTION: GLUCURONIDE REPRESSORS AND USES THEREOF
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:

; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/882,704A
; FILING DATE: 25-JUN-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5879906tenburg Ph.D., Carol
; REGISTRATION NUMBER: 39,317
; REFERENCE/DOCKET NUMBER: 190106.404
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 457 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-882-704A-6

Query Match 6.8%; Score 81.5; DB 1; Length 457;
Best Local Similarity 25.2%; Pred. No. 2.9;
Matches 37; Conservative 30; Mismatches 53; Indels 27; Gaps 8;

QY 95 WDLRSGLGAVFSRVNNV-----VLEA-----PFLVGIEGSLKSTYNLL-FCGSCG 141
Db 317 WSLPVALVALAIAASIGQGVTTVMWALEADTVEYGYLTGVR--IEGLTYSLSFSTRKCG 374

QY 142 IPVGPHLYSTHAALALRGHFC--LSSDKMVCYLLKTKAIVNASMDIQNV-----PLSE 194
Db 375 QAIG--GSIPAFILGSLGYIANQVTPVINGIRTSIALVPCGFMLLAFVIWFPYPLTD 431

QY 195 KIAELKEKIVLTHNRKLSMKILSEVT 221
Db 432 K--KFEIVVEIDNRKKVQQQLISDIT 456

RESULT 5
US-09-151-957-6
; Sequence 6, Application US/09151957
; Patent No. 6429292
; GENERAL INFORMATION:
; APPLICANT: Jefferson, Richard A.
; Wilson, Katherine J.
; Leader, Michael
; TITLE OF INVENTION: GLUCURONIDE REPRESSORS AND USES THEREOF
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/151,957
; FILING DATE: 11-Sep-1998
; CLASSIFICATION: <Unknown>

Db 134 LISASGQLKIADFGFLARVFS---PDGSRILY-THQ--VATRSVGCIMGELLNGSLPLFPQGN 187
Qy 168 --KWCYLLKTKAIIVNASEM-DIQNVPLSEKIAELKEKIVLTHNRLKSLMKILSVTPDQ 224
Db 188 DIEQLCYVLRILGTNPQWPPELTLPDYNKIS-FKEQVPM-----LEEVLDPV 236
Qy 225 S 225
Db 237 S 237

RESULT 8

US-09-018-576-12
; Sequence 12, Application US/09018576
; Patent No. 5968800
; GENERAL INFORMATION:
; APPLICANT: Gerhold, David L.
; TITLE OF INVENTION: CYCLIN-DEPENDENT PROTEIN KINASE
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: P.O. Box 2000, RY60-30
; CITY: Rahway
; STATE: NJ
; COUNTRY: US
; ZIP: 07065-0907
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/018,576
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Hand, J. Mark
; REGISTRATION NUMBER: 36,545
; REFERENCE/DOCKET NUMBER: 19885Y
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 732/594-3905
; TELEFAX: 732/594-4720
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 325 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-018-576-12

Query Match 6.6%; Score 78.5; DB 1; Length 325;
Best Local Similarity 23.2%; Pred. No. 3.8;
Matches 56; Conservative 24; Mismatches 78; Indels 83; Gaps 12;
Qy 22 GGTERRAIDQASFTTSMEDTQVVGSSPLGPAGLG-----AEEPAAG 63
Db 43 GFNPQALREIKALQEME-DNQYVQLKAVFPHGGLVLAPEFMLSALAEVVRHQRPLAQ 101
Qy 64 POLPSWLOPERCAVFCQACQACHAVLADSLRSLGAVVFSRVTVNNV-----LEAPF 119
Db 102 AQVKSILQMLKGV---AFCHA-----NNIVHRLKPNAL 133
Qy 120 LVGEGSLKSTYLLFCGSCGIPVGFHLYSTHAALALRGHFCSSD----- 167
Db 134 LISASGQLKIADFGFLARVFS---PDGSRILY-THQ--VATRSVGCIMGELLNGSLPLFPQGN 187
Qy 168 --KWCYLLKTKAIIVNASEM-DIQNVPLSEKIAELKEKIVLTHNRLKSLMKILSVTPDQ 224
Db 188 DIEQLCYVLRILGTNPQWPPELTLPDYNKIS-FKEQVPM-----LEEVLDPV 236
Qy 225 S 225

Db 237 S 237
RESULT 9
US-09-248-137-3
; Sequence 3, Application US/09248137
; Patent No. 6030788
; GENERAL INFORMATION:
; APPLICANT: Gerhold, David L.
; TITLE OF INVENTION: CYCLIN-DEPENDENT PROTEIN KINASE
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: P.O. Box 2000, RY60-30
; CITY: Rahway
; STATE: NJ
; COUNTRY: US
; ZIP: 07065-0907
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/248,137
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/018,576
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Hand, J. Mark
; REGISTRATION NUMBER: 36,545
; REFERENCE/DOCKET NUMBER: 19885Y
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 732/594-3905
; TELEFAX: 732/594-4720
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 325 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-248-137-3

Query Match 6.6%; Score 78.5; DB 2; Length 325;
Best Local Similarity 23.2%; Pred. No. 3.8;
Matches 56; Conservative 24; Mismatches 78; Indels 83; Gaps 12;
Qy 22 GGTERRAIDQASFTTSMEDTQVVGSSPLGPAGLG-----AEEPAAG 63
Db 43 GFNPQALREIKALQEME-DNQYVQLKAVFPHGGLVLAPEFMLSALAEVVRHQRPLAQ 101
Qy 64 POLPSWLOPERCAVFCQACQACHAVLADSLRSLGAVVFSRVTVNNV-----LEAPF 119
Db 102 AQVKSILQMLKGV---AFCHA-----NNIVHRLKPNAL 133
Qy 120 LVGEGSLKSTYLLFCGSCGIPVGFHLYSTHAALALRGHFCSSD----- 167
Db 134 LISASGQLKIADFGFLARVFS---PDGSRILY-THQ--VATRSVGCIMGELLNGSLPLFPQGN 187
Qy 168 --KWCYLLKTKAIIVNASEM-DIQNVPLSEKIAELKEKIVLTHNRLKSLMKILSVTPDQ 224
Db 188 DIEQLCYVLRILGTNPQWPPELTLPDYNKIS-FKEQVPM-----LEEVLDPV 236
Qy 225 S 225
Db 237 S 237

RESULT 10

US-09-248-137-12
; Sequence 12, Application US/09248137
; Patent No. 6030788
; GENERAL INFORMATION:
; APPLICANT: Gerhold, David L.
; TITLE OF INVENTION: CYCLIN-DEPENDENT PROTEIN KINASE
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: P.O. Box 2000, RY60-30
; CITY: Rahway
; STATE: NJ
; COUNTRY: US
; ZIP: 07065-0907
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/248,137
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/018,576
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Hand, J. Mark
; REGISTRATION NUMBER: 36,545
; REFERENCE/DOCKET NUMBER: 19885Y
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 732/594-3905
; TELEFAX: 732/594-4720
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 325 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-248-137-12

Query Match 6.6%; Score 78.5; DB 2; Length 325;
Best Local Similarity 23.2%; Pred. No. 3.8; Indels 83; Gaps 12;
Matches 56; Conservative 24; Mismatches 78; Indels 83; Gaps 12;
QY 22 GGERAIDQASFTTSMENDTQVVGSSPLGPAGLG-----AEEPAAG 63
Db 43 GPPNALREIKALQEME-DNQVYVQLKAVFPHGGGVFLAFEFMLSDLAERVVHQAORPLAQ 101
QY 64 POLPSWLQPERCAVFCQACQACHAVLADSVHLAWDLRSLSGAVVFSRVTVNNV-----LEAPF 119
Db 102 AQVKSYLQMLLKGKVA---AFCHA-----NNIVHRNLKPNAL 133
QY 120 LVGIEGSLKGYTNLLFCGSCGIPVGFHLYSTHAAALRGHFCSSD----- 167
Db 134 LISASGQLKIADFGARVFS---PGSRLY-THQ--VATRSVGCINGELLNGSLPFGKN 187
QY 168 --KMVCYLLKTKKAIIVNASEM-DIQNVPLSEKIAELKEKIVLTHNRLKSLMKILSEVTPDQ 224
Db 188 DIEQLCYVLRILGTNPQWPWPELTLPDYNKIS-FKEQVMP-----LEEVLFDV 236
QY 225 S 225
Db 237 S 237

RESULT 11
US-09-771-161A-237
; Sequence 237, Application US/09771161A
; Patent No. 6936450
; GENERAL INFORMATION:
; APPLICANT: LEVINE, et al.

; TITLE OF INVENTION: VARIANTS OF PROTEIN KINASES
; FILE REFERENCE: 802620-2005.1
; CURRENT APPLICATION NUMBER: US/09/771,161A
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 09/724,676
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 136776
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 135619
; PRIOR FILING DATE: 2000-04-12
; NUMBER OF SEQ ID NOS: 273
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 237
; LENGTH: 452
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-771-161A-237

Query Match 6.6%; Score 78.5; DB 2; Length 452;
Best Local Similarity 23.2%; Pred. No. 6.3; Indels 83; Gaps 12;
Matches 56; Conservative 24; Mismatches 78; Indels 83; Gaps 12;
QY 22 GGERAIDQASFTTSMENDTQVVGSSPLGPAGLG-----AEEPAAG 63
Db 43 GPPNALREIKALQEME-DNQVYVQLKAVFPHGGGVFLAFEFMLSDLAERVVHQAORPLAQ 101
QY 64 POLPSWLQPERCAVFCQACQACHAVLADSVHLAWDLRSLSGAVVFSRVTVNNV-----LEAPF 119
Db 102 AQVKSYLQMLLKGKVA---AFCHA-----NNIVHRNLKPNAL 133
QY 120 LVGIEGSLKGYTNLLFCGSCGIPVGFHLYSTHAAALRGHFCSSD----- 167
Db 134 LISASGQLKIADFGARVFS---PGSRLY-THQ--VATRSVGCINGELLNGSLPFGKN 187
QY 168 --KMVCYLLKTKKAIIVNASEM-DIQNVPLSEKIAELKEKIVLTHNRLKSLMKILSEVTPDQ 224
Db 188 DIEQLCYVLRILGTNPQWPWPELTLPDYNKIS-FKEQVMP-----LEEVLFDV 236
QY 225 S 225
Db 237 S 237

RESULT 12
US-09-252-991A-19049
; Sequence 19049, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 19049
; LENGTH: 190
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19049

Query Match 6.5%; Score 78; DB 2; Length 190;
Best Local Similarity 32.9%; Pred. No. 1.9; Indels 14; Gaps 3;
Matches 24; Conservative 6; Mismatches 29; Indels 14; Gaps 3;
QY 4 QPLRHRRCATPPRGDFCGGTERAIDQASFTTSMENDTQVVGSSPLGPAGLGAEEPAAG 63
Db 94 RPCRHRRCVTPAAGGYPG---RAGDRAS-----VARSTGPGLPAAAPGRPRGHC- 139

QY 64 POLPSWQPERCA 76
DB 140 PAPPWPRRRCA 152

RESULT 13
US-09-537-682-1
; Sequence 1, Application US/09537682
; Patent No. 6303357
; GENERAL INFORMATION:
; APPLICANT: TAKUCHI, Kenichi
; APPLICANT: KOIDE, Yoshinao
; APPLICANT: NAKAMISHI, Yuji
; APPLICANT: SUZUKI, Satoru
; TITLE OF INVENTION: L-ALPHA-GLYCEROPHOSPHATE OXIDASE GENE, RECOMBINANT
; TITLE OF INVENTION: DNA, AND METHOD FOR PRODUCING MODIFIED
; FILE REFERENCE: A20-121814C/KI
; CURRENT FILING DATE: 2000-03-29
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 1
; TYPE: PRT
; LENGTH: 607
; ORGANISM: Enterococcus faecium No. 6303357 7044
US-09-537-682-1

Query Match 6.4%; Score 77; DB 2; Length 607;
Best Local Similarity 23.2%; Pred. No. 15;
Matches 49; Conservative 30; Mismatches 76; Indels 56; Gaps 10;

QY 19 DFCGTERAIDQASFTSM-----EWDQVVGKSSPLGP-----AGLGAEPPAAGPQ 65
DB 51 DPAEFT-----SSRSTKLHVGIGIRYKTFDVEW--ADTVGERAVVQGIAPHIPKPDPM 102
QY 66 LPSWQPERCAVQCAQCHAVLADSVHLAWDLRSRSGAVVFSRVN-----NVLEAPP 119
DB 103 LPIIYEDGATTFNMP-----SVKVAMDLYDKLANVTGKYENTLTTPBEVLREPP 154
QY 120 LVGIGSLKGYNYLLFCGSGGIPGVF-----HLYSTHAALALRGHFCLSDDKWCYLL 174
DB 155 LK--KEGLKA-----GYLDFRNDARLVLDNIKKAAEDGAYLVSKWKAAGFLY 202
QY 175 KTKAIVNAEMDIQNVPLSEKIAELKEKIVL 205
DB 203 EGDQIVGVKARDL-----LTDEVIEIKSLVI 229

RESULT 14
US-08-520-933-3
; Sequence 3, Application US/08520933
; Patent No. 5981194
; GENERAL INFORMATION:
; APPLICANT: Jefferies, Wilfred A.
; APPLICANT: McGeer, Patrick L.
; APPLICANT: Rothenberger, Sylvia
; APPLICANT: Food, Michael R.
; APPLICANT: Yamada, Tatsuo
; APPLICANT: Kennard, Malcolm
; TITLE OF INVENTION: Use of p97 and Iron Binding Proteins
; TITLE OF INVENTION: as Diagnostic and Therapeutic Agents
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bereskin & Parr
; STREET: 40 King Street West
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5H 3Y2
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.25

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/520,933
; FILING DATE: August 31, 1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Shona S. McDiarmid
; REGISTRATION NUMBER: 38,798
; REFERENCE/DOCKET NUMBER: 7685-006
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 416-364-7311
; TELEFAX: 416-361-1398
; TELEX: 06-23115
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 719 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-520-933-3

Query Match 6.4%; Score 77; DB 1; Length 719;
Best Local Similarity 23.7%; Pred. No. 19;
Matches 59; Conservative 25; Mismatches 87; Indels 78; Gaps 15;

QY 17 RGDGCGTERAIDQASFTSM-----DTQVVGKSSPLGPAGLGAEEPAAGP 64
DB 171 RGDSSG--EGVCDKSPLERYDYDGAFCRLAEGAGDVAFVGHSTVL-----ENTDGK 220
QY 65 QLPWQLP-----ERCAVFOCAQCH--AVLADSVHLAWDLRSRSGAVVFSRV 109
DB 221 TLPFGQALLSODFELLCRDSRADVTWRQCHLARVPAHVVVRADTD---GGLIF-RL 276
QY 110 TNNVLEAPFLVGIERS-----LKGSTY---NLLFCGSCG--IPVGFHLYST---HAALAA 157
DB 277 LN-----EQRLFSHSGSFQPFSEAYCGKOLLFKDSTSELVPIATQTYEAMLGHEYLHA 332
QY 158 LRGHFCLSDKWCYLL-----LTKAIVNAEMDIQNVPLSEKIAELKEKIVLTHNRLKSL 213
DB 333 MKGLLC-DPNRLPPVLRWCVLSTPBIQKGDH-----AVAFRRQRLKPE 375
QY 214 MKLSEVTP 222
DB 376 IQCVSAKSP 384

RESULT 15
US-09-285-040-3
; Sequence 3, Application US/09285040
; Patent No. 6455494
; GENERAL INFORMATION:
; APPLICANT: Jefferies, Wilfred A.
; APPLICANT: McGeer, Patrick L.
; APPLICANT: Rothenberger, Sylvia
; APPLICANT: Food, Michael R.
; APPLICANT: Yamada, Tatsuo
; APPLICANT: Kennard, Malcolm
; TITLE OF INVENTION: Use of p97 and Iron Binding Proteins
; TITLE OF INVENTION: as Diagnostic and Therapeutic Agents
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bereskin & Parr
; STREET: 40 King Street West
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5H 3Y2
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25

Search completed: March 25, 2006, 08:04:49
Job time : 48 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 25, 2006, 08:14:45 ; Search time 165 Seconds
(without alignments)
579.896 Million cell updates/sec

Title: US-09-942-052A-728
Perfect score: 1198
Sequence: 1 MAAQPLRHRSRCATPRGDF.....LKSLMKILSEVTPDQSKPEN 229

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA Main:

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1198	100.0	229	3	US-09-942-052-728
2	1198	100.0	229	3	US-09-942-052-728
3	1198	100.0	229	3	US-09-942-052-730
4	1198	100.0	229	3	US-09-942-052-731
5	138.5	11.6	164	3	US-09-942-052-707
6	138	11.5	29	3	US-09-942-052-704
7	113	9.4	233	3	US-09-890-688-82
8	113	9.4	233	4	US-10-408-765A-969
9	95.5	8.0	708	4	US-10-437-963-195427
10	92	7.7	361	4	US-10-437-963-154548
11	92	7.7	695	4	US-10-369-493-18389
12	89	7.4	664	3	US-09-780-525-2
13	88.5	7.4	435	4	US-10-437-963-186569
14	87.5	7.3	848	4	US-10-437-963-155606
15	86.5	7.2	669	4	US-10-282-122A-51802
16	85	7.1	337	4	US-10-156-761-10954
17	84.5	7.1	485	4	US-10-363-829-358
18	84	7.0	295	4	US-10-424-599-223506
19	84	7.0	1902	4	US-10-437-963-139559
20	83.5	7.0	888	4	US-10-437-963-147897
21	83	6.9	663	4	US-10-282-122A-51673
22	82.5	6.9	124	4	US-10-425-115-216789
23	82.5	6.9	278	4	US-10-425-114-42310
24	82.5	6.9	325	4	US-10-663-896-2
25	82.5	6.9	493	5	US-10-213-974-36
26	82.5	6.9	502	4	US-10-424-599-226730
27	82.5	6.9	1043	4	US-10-369-493-9793

28 81.5 6.8 278 4 US-10-425-114-66220 Sequence 66220, A
29 81.5 6.8 457 4 US-10-195-518-6 Sequence 6, Appli
30 81.5 6.8 701 5 US-10-450-763-31748 Sequence 31748, A
31 81.5 6.8 1969 4 US-10-437-963-115342 Sequence 115342,
32 81.5 6.8 3067 5 US-10-631-467-1618 Sequence 1618, Ap
33 81 6.8 229 4 US-10-425-115-281194 Sequence 281194,
34 81 6.8 859 4 US-10-210-281-60 Sequence 60, Appl
35 80.5 6.7 282 5 US-10-450-763-51058 Sequence 51058, A
36 80.5 6.7 1711 5 US-10-029-345A-38 Sequence 38, Appl
37 80 6.7 305 4 US-10-183-116-59 Sequence 59, Appl
38 80 6.7 305 5 US-10-957-135-59 Sequence 59, Appl
39 80 6.7 305 6 US-11-083-611-59 Sequence 158084,
40 80 6.7 544 4 US-10-424-599-158084 Sequence 968, App
41 79.5 6.6 212 3 US-09-925-299-968 Sequence 968, App
42 79.5 6.6 212 3 US-09-925-299-968 Sequence 57058, A
43 79 6.6 196 5 US-10-450-763-44924 Sequence 44924, A
44 79 6.6 230 5 US-10-450-763-44924 Sequence 44919, A
45 79 6.6 271 5 US-10-450-763-44919

ALIGNMENTS

RESULT 1
US-09-942-052-728
; Sequence 728, Application US/09942052
; Publication NO. US20030170626A1
; GENERAL INFORMATION:
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Paris, Mary
; APPLICANT: Hubert, Rene S.
; APPLICANT: Afar, Daniel
; APPLICANT: Ge, Wangmao
; APPLICANT: Challita-Eid, Pia M.
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN ENTITLED 85P1B3
; FILE REFERENCE: 51158-20028.00
; CURRENT APPLICATION NUMBER: US/09/942,052
; CURRENT FILING DATE: 2001-08-28
; PRIOR APPLICATION NUMBER: 60/228,432
; PRIOR FILING DATE: 2000-08-28
; NUMBER OF SEQ ID NOS: 744
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 728
; LENGTH: 229
; TYPE: PRT
; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: 85P1B3/OIP5
; OTHER INFORMATION: Clone A protein
US-09-942-052-728

Query Match 100.0%; Score 1198; DB 3; Length 229;
Best Local Similarity 100.0%; Pred. No. 9.8e-114;
Matches 229; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAAQPLRHRSRCATPRGDFCGGTERAIDQASFTTSMEDTQVKGSSPLGPAAGLGAERP 60
Db 1 MAAQPLRHRSRCATPRGDFCGGTERAIDQASFTTSMEDTQVKGSSPLGPAAGLGAERP 60
QY 61 AAGPOLPSWLOPERCAVFOCAQCHAVLADSVHLAWDLRSGLGAVVFSRVTVNNVLEAPPL 120
Db 61 AAGPOLPSWLOPERCAVFOCAQCHAVLADSVHLAWDLRSGLGAVVFSRVTVNNVLEAPPL 120
QY 121 VGIEGSLKGYTYNLLFCGSCCIPVGFHYLSTHAALALRGHFCGLSSDRKWCYLLKTKAIV 180
Db 121 VGIEGSLKGYTYNLLFCGSCCIPVGFHYLSTHAALALRGHFCGLSSDRKWCYLLKTKAIV 180
QY 181 NASEMDIQNVPLSKIAELKEKIVLTHNRLKSLMKILSEVTPDQSKPEN 229
Db 181 NASEMDIQNVPLSKIAELKEKIVLTHNRLKSLMKILSEVTPDQSKPEN 229


```
RESULT 2
US-09-942-052-729
; Sequence 729, Application US/09942052
; Publication No. US20030170626A1
; GENERAL INFORMATION:
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Paris, Mary
; APPLICANT: Hubert, Rene S.
; APPLICANT: Afar, Daniel
; APPLICANT: Ge, Wangmao
; APPLICANT: Challita-Bid, Pia M.
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN ENTITLED 85P1B3
; TITLE OF INVENTION: USEFUL IN TREATMENT AND DETECTION OF CANCER
; FILE REFERENCE: 51158-20028.00
; CURRENT APPLICATION NUMBER: US/09/942,052
; CURRENT FILING DATE: 2001-08-28
; PRIOR APPLICATION NUMBER: 60/228,432
; PRIOR FILING DATE: 2000-08-28
; NUMBER OF SEQ ID NOS: 744
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 729
; LENGTH: 229
; TYPE: PRT
; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: 85P1B3/OIP5
; OTHER INFORMATION: protein sequence
US-09-942-052-729

Query Match 100.0%; Score 1198; DB 3; Length 229;
Best Local Similarity 100.0%; Pred. No. 9.8e-114;
Matches 229; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAAQPLRHRSRCATPPRGDFCGGTERAIDQASFTTSMEDTQVVKGSPLGPAGLGAEPP 60
Db 1 MAAQPLRHRSRCATPPRGDFCGGTERAIDQASFTTSMEDTQVVKGSPLGPAGLGAEPP 60

Qy 61 AAGPOLPSWLQPERCAVFOCAQCHAVLADSVHLAWDLRSLSGAVVFSRVTVNNVLEAPFL 120
Db 61 AAGPOLPSWLQPERCAVFOCAQCHAVLADSVHLAWDLRSLSGAVVFSRVTVNNVLEAPFL 120

Qy 121 VGIEGSLKGYSTYLLFCGSCGIPVGFHYLSTHAALALRGHFCILSSDKMVCYLLKTKAIV 180
Db 121 VGIEGSLKGYSTYLLFCGSCGIPVGFHYLSTHAALALRGHFCILSSDKMVCYLLKTKAIV 180

Qy 181 NASEMDIQNVPLSEKIAELKEKIVLTHNRLKSLMKILSEVTPDQSKPEN 229
Db 181 NASEMDIQNVPLSEKIAELKEKIVLTHNRLKSLMKILSEVTPDQSKPEN 229

RESULT 3
US-09-942-052-730
; Sequence 730, Application US/09942052
; Publication No. US20030170626A1
; GENERAL INFORMATION:
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Paris, Mary
; APPLICANT: Hubert, Rene S.
; APPLICANT: Afar, Daniel
; APPLICANT: Ge, Wangmao
; APPLICANT: Challita-Bid, Pia M.
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN ENTITLED 85P1B3
; TITLE OF INVENTION: USEFUL IN TREATMENT AND DETECTION OF CANCER
; FILE REFERENCE: 51158-20028.00
; CURRENT APPLICATION NUMBER: US/09/942,052
; CURRENT FILING DATE: 2001-08-28
; PRIOR APPLICATION NUMBER: 60/228,432
; PRIOR FILING DATE: 2000-08-28
; NUMBER OF SEQ ID NOS: 744
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 730
; LENGTH: 229
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-942-052-731

Query Match 100.0%; Score 1198; DB 3; Length 229;
Best Local Similarity 100.0%; Pred. No. 9.8e-114;
Matches 229; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAAQPLRHRSRCATPPRGDFCGGTERAIDQASFTTSMEDTQVVKGSPLGPAGLGAEPP 60
Db 1 MAAQPLRHRSRCATPPRGDFCGGTERAIDQASFTTSMEDTQVVKGSPLGPAGLGAEPP 60

Qy 61 AAGPOLPSWLQPERCAVFOCAQCHAVLADSVHLAWDLRSLSGAVVFSRVTVNNVLEAPFL 120
Db 61 AAGPOLPSWLQPERCAVFOCAQCHAVLADSVHLAWDLRSLSGAVVFSRVTVNNVLEAPFL 120

Qy 121 VGIEGSLKGYSTYLLFCGSCGIPVGFHYLSTHAALALRGHFCILSSDKMVCYLLKTKAIV 180
Db 121 VGIEGSLKGYSTYLLFCGSCGIPVGFHYLSTHAALALRGHFCILSSDKMVCYLLKTKAIV 180

Qy 181 NASEMDIQNVPLSEKIAELKEKIVLTHNRLKSLMKILSEVTPDQSKPEN 229
Db 181 NASEMDIQNVPLSEKIAELKEKIVLTHNRLKSLMKILSEVTPDQSKPEN 229

RESULT 4
US-09-942-052-731
; Sequence 731, Application US/09942052
; Publication No. US20030170626A1
; GENERAL INFORMATION:
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Paris, Mary
; APPLICANT: Hubert, Rene S.
; APPLICANT: Afar, Daniel
; APPLICANT: Ge, Wangmao
; APPLICANT: Challita-Bid, Pia M.
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN ENTITLED 85P1B3
; TITLE OF INVENTION: USEFUL IN TREATMENT AND DETECTION OF CANCER
; FILE REFERENCE: 51158-20028.00
; CURRENT APPLICATION NUMBER: US/09/942,052
; CURRENT FILING DATE: 2001-08-28
; PRIOR APPLICATION NUMBER: 60/228,432
; PRIOR FILING DATE: 2000-08-28
; NUMBER OF SEQ ID NOS: 744
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 731
; LENGTH: 229
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-942-052-731

Query Match 100.0%; Score 1198; DB 3; Length 229;
Best Local Similarity 100.0%; Pred. No. 9.8e-114;
Matches 229; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAAQPLRHRSRCATPPRGDFCGGTERAIDQASFTTSMEDTQVVKGSPLGPAGLGAEPP 60
Db 1 MAAQPLRHRSRCATPPRGDFCGGTERAIDQASFTTSMEDTQVVKGSPLGPAGLGAEPP 60

Qy 61 AAGPOLPSWLQPERCAVFOCAQCHAVLADSVHLAWDLRSLSGAVVFSRVTVNNVLEAPFL 120
Db 61 AAGPOLPSWLQPERCAVFOCAQCHAVLADSVHLAWDLRSLSGAVVFSRVTVNNVLEAPFL 120

Qy 121 VGIEGSLKGYSTYLLFCGSCGIPVGFHYLSTHAALALRGHFCILSSDKMVCYLLKTKAIV 180
Db 121 VGIEGSLKGYSTYLLFCGSCGIPVGFHYLSTHAALALRGHFCILSSDKMVCYLLKTKAIV 180

Qy 181 NASEMDIQNVPLSEKIAELKEKIVLTHNRLKSLMKILSEVTPDQSKPEN 229
Db 181 NASEMDIQNVPLSEKIAELKEKIVLTHNRLKSLMKILSEVTPDQSKPEN 229
```


Qy 208 NLRKSLMKILSE 219
Db 210 DVLKALQMKLWE 221

RESULT 8

US-10-408-765A-969

; Sequence 969, Application US/10408765A

; Publication No. US20040101874A1

; GENERAL INFORMATION:

; APPLICANT: Ghosh, Soumitra S.

; APPLICANT: Faby, Eoin D.

; APPLICANT: Zhang, Bing

; APPLICANT: Gibson, Bradford W.

; APPLICANT: Taylor, Steven W.

; APPLICANT: Glenn, Gary M.

; APPLICANT: Warnock, Dale E.

; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION

; FILE REFERENCE: 660088.465

; CURRENT APPLICATION NUMBER: US/10/408, 765A

; CURRENT FILING DATE: 2003-04-04

; NUMBER OF SEQ ID NOS: 3077

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 969

; LENGTH: 233

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-408-765A-969

Query Match 9.4%; Score 113; DB 4; Length 233;

Best Local Similarity 27.1%; Pred. No. 0.0075;

Matches 52; Conservative 30; Mismatches 88; Indels 22; Gaps 8;

Qy 31 ASFTSMEDTQVVKSSPLGPAAGAEPAAGPOLPSWLOPERCAVFOCAOCHAVLADS 90

Db 49 ASWSSMSDASV-----ADMERQL--EEEAARAE-----EPVFLSCGCRPLGDS 95

Qy 91 VHLAWLGR--SLGAVFVRVNNVLEAPPLVIGESLKGSTYNNLLFCGSGCIPVGHLY 149

Db 96 --LSWVASQEDTNCILLRCVSCNVSDVDEKQKSKREKENGCVLETLCCAGCSLNLGVYR 153

Qy 150 STHAALAAHGFCLSSDKMKVCYL--KTKAIVNASEMDIONVPLSEKIAELKEKIVLTH 207

Db 154 CTPLNDYKRDPLCLSVESVIGSSSEKQIV--SEDKELFNL---ESRVEIKSLTQME 209

Qy 208 NLRKSLMKILSE 219

Db 210 DVLKALQMKLWE 221

RESULT 9

US-10-437-963-195427

; Sequence 195427, Application US/10437963

; Publication No. US20040123343A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.

; APPLICANT: Kovalic, David K.

; APPLICANT: Zhou, Yihua

; APPLICANT: Cao, Yongwei

; APPLICANT: Wu, Wei

; APPLICANT: Boukharov, Andrey A.

; APPLICANT: Barbazuk, Brad

; APPLICANT: Li, Ping

; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With

; FILE REFERENCE: 38-21(53221)B

; CURRENT APPLICATION NUMBER: US/10/437, 963

; CURRENT FILING DATE: 2003-05-14

; NUMBER OF SEQ ID NOS: 204966

; SEQ ID NO 195427

; LENGTH: 708

; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_91377C.1.pep
US-10-437-963-195427

Query Match 8.0%; Score 95.5; DB 4; Length 708;

Best Local Similarity 24.5%; Pred. No. 2.1;

Matches 71; Conservative 32; Mismatches 86; Indels 101; Gaps 17;

Qy 6 LRHRSRCATPPRGDF-----CGGTERAIDQASFTTSMEW 39

Db 225 LRHRKANTESASDASPPPRQSIIVTIGKEARAKAARAKSGTSA-----SPTTVST 279

Qy 40 DTQVVKSSPLGPAAGAEPAAGPOLP-----SW--LOPERCAVF-----OCAQCH 84

Db 280 DVVVFVVGSGQVTPSG--PISDPAGGSLPEAVLTWEEQLQVEMGRLLKAGAGIGREISEAR 338

Qy 85 AVLADSVHLAWDLRSISG-----AVFVRVTVNNVLE-----APLVGIEGSLK 128

Db 339 AETAAANALAERLVRELAEREDELTKMRELVAGNEROSKLEDRMSLGNLSEIRGSLR 398

Qy 129 GSTYNLL--FCGSCG-----IPVG---PHLYSTHAALALRG-----HFCLSSDKM----- 169

Db 399 -VYTGHLQLAKECGVKSTIPVNLDFSLTSSLAELATANGEPISKHTSRIAEETNGIY 457

Qy 170 --VCYLLKTKAIVNAS--EMDIONV-----PLSEKIAELKEKIV 204

Db 458 TGVCHVL---ACVRLSRPELDLREILDQGAASDTRKEVMEVSDLSGSLV 504

RESULT 10

US-10-437-963-154548

; Sequence 154548, Application US/10437963

; Publication No. US20040123343A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.

; APPLICANT: Kovalic, David K.

; APPLICANT: Zhou, Yihua

; APPLICANT: Cao, Yongwei

; APPLICANT: Wu, Wei

; APPLICANT: Boukharov, Andrey A.

; APPLICANT: Barbazuk, Brad

; APPLICANT: Li, Ping

; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With

; FILE REFERENCE: 38-21(53221)B

; CURRENT APPLICATION NUMBER: US/10/437, 963

; CURRENT FILING DATE: 2003-05-14

; NUMBER OF SEQ ID NOS: 204966

; SEQ ID NO 154548

; LENGTH: 361

; TYPE: PRT

; ORGANISM: Oryza sativa

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT_MRT4530_54398C.1.pep

US-10-437-963-154548

Query Match

Best Local Similarity 25.3%; Pred. No. 1.9;

Matches 61; Conservative 28; Mismatches 92; Indels 60; Gaps 14;

Qy 2 AAQPLHRSRCATPPRGDFCGGTERAIDQASFTTSMEDTQVVKSSPLGPAAGAEPA 61

Db 49 APTPLFRRAVRAKAAQDGGSGTSSAPAVAST-----DVVVVPGSREATPFG--PASDPV 102

Qy 62 AGPOLP-----SW--LOPERCAVFOCAOCHAVLADSVHLAWDLRSGLGAVFVRVTVNNVVL 115

Db 103 AGRGSPAIVLSWEELQVEMGRLLLEACA--RVIGRETAERGLEHRM-----SELGNN--- 152

Qy 116 EAPFLVIGIEGSLKGSTYNLL--FCGSCGI-----PVGPHLYSTHAALAL-----RGH 161

Db 153 ----LSEIRGSLR-VYTTGLHQLAGCKGIKSTIPANPDEFSLTSSLAELAAAMEEIPSKH 207

QY 162 FCLSDKM-----VCYLLTKTKAIVNASEMDIQNV-----PLSEKIAELKXKI 203
Db 208 AARIGEEMSNRYIGACHILACVRLAH-PELDREILDQGEASDARKDQVMEVGDGKSV 266
QY 204 V 204
Db 267 L 267

RESULT 11

US-10-369-493-18389
; Sequence 18389, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 18389
; LENGTH: 695
; TYPE: PRT
; ORGANISM: Lactococcus lactis
US-10-369-493-18389

Query Match 7.7%; Score 92; DB 4; Length 695;
Best Local Similarity 25.3%; Pred. No. 4.7;
Matches 38; Conservative 28; Mismatches 50; Indels 34; Gaps 7;
QY 101 LGAVVSRVTVNNVLEAPFLVIGIEGSLKSTYNLLFCGSCGIPVGFHLYSTHAAALRG 160
Db 101 LGTIIP-----FYSGGTFFPSGARGELKSRKPMMLITWGITVAY-AYSVIATIMSLG 153
QY 161 HF-----CLSSDKMVCYLLTKAIVNASE-MDI-----QNVPLSE-K 195
Db 154 HMGNFVFEATLIVIMLGHLEIMKAIMGAGDALDOLASLVKPKAHLKSGKDVLSLK 213
QY 196 IAEKKEKIVLTHRLKSLMKILSEVTPDQS 225
Db 214 VGDL--LLVKNENKIPADGILSEALVDES 241

RESULT 12

US-09-780-525-2
; Sequence 2, Application US/09780525
; Patent No. US20020004223A1
; GENERAL INFORMATION:
; APPLICANT: Bin-Bing Zhou
; APPLICANT: Yuan Zhu
; APPLICANT: Priya Chaturvedi
; APPLICANT: Mark R. Hurle
; APPLICANT: Xiaotong Li
; TITLE OF INVENTION: PHARI, A NEW RING FINGER PROTEIN
; FILE REFERENCE: GP-70668-C1
; CURRENT APPLICATION NUMBER: US/09/780,525
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 09/456,876
; PRIOR FILING DATE: 1999-12-08
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 664
; TYPE: PRT
; ORGANISM: HOMO SAPIENS

US-09-780-525-2

Query Match 7.4%; Score 89; DB 3; Length 664;
Best Local Similarity 23.0%; Pred. No. 8.9;
Matches 52; Conservative 26; Mismatches 74; Indels 74; Gaps 11;
QY 8 HRSRCATPPRGDFC-----CGTERAIDQA-----SFTTSMENDTVQVKGSSPL----- 50
Db 434 YRQQAQPPH---CPAPEGEPGAPQALGDAPSTSVSLTAVQDYVCPQLQGSHALCTCCFP 490
QY 51 ---GPAGLGAEEPAAGPOLPSWLPQRCVAVFQCAQCHAVLADSVHLANDLSR----- 99
Db 491 PMPDRAERQDPRVAQ-----QCAVC-----LQPFCHLYWGCTRTGCGCLA 534
QY 100 -----SLGAVVSRVTVNNVLEAPFLVIGIEGSLKSTY-NLLFCGSCGIPVGFHLYSTHA 153
Db 535 PFCNLNLGDKLGDGVNNSVESDILKNYLAT-RGLTWKML-----TES 578
QY 154 ALAALRGHFCLS-----SDKMVCYLLTKAIVNASEMDIQNVPLSE 194
Db 579 LVALLQRGVFLLSRVYRTGDTVLVCYCCGLRSFRELTYQYRONIPASE 624

RESULT 13

US-10-437-963-186569
; Sequence 186569, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 186569
; LENGTH: 435
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_83357C.1.pep
US-10-437-963-186569

Query Match 7.4%; Score 88.5; DB 4; Length 435;
Best Local Similarity 22.9%; Pred. No. 5.6;
Matches 50; Conservative 27; Mismatches 82; Indels 59; Gaps 12;
QY 2 AAQPL--RH--RSRCATPPRGDFCGGTERAIDQAFTTSMENDTVQVKGSSPLGPAGLGA 57
Db 41 APEPLSCRHGRHLKCAVD-----GGAGRETERPSPAPQ-----RESPSGSLGAAL 88
QY 58 EEPAGPQLPSWLPQRCVAVFQCAQCHA--VLADSVHLAWDLRSLSGAVVSRVTVNNVL 115
Db 89 EDPSPQPGVPLILLP-----LCCRCYAKEICSEYVVRITDLVNI-----LNSNAIS 135
QY 116 EAPFLVIGIEGSLKSTYNLLFCGSCGIPVGFHLYSTHAAALRGHFCLSDDKVCYLLK 175
Db 136 EGPFF-----SMRKARFLGSAS---AFSVKQTEWP-----CATTSKKCY-LQ 173
QY 176 TKAIVNASEMDIQN-----VPLSEKIAELKXKIVLTH 207
Db 174 NGSFGGITEDEQSSLYNFLYPSKELLPPDKEMSIIDH 211

RESULT 14

US-10-437-963-155606

```

; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 51802
; LENGTH: 669
; TYPE: PRT
; ORGANISM: Clostridium acetobutylicum
US-10-282-122A-51802

Query Match 7.2%; Score 86.5; DB 4; Length 669;
Best Local Similarity 27.4%; Pred. No.16;
Matches 37; Conservative 22; Mismatches 45; Indels 31; Gaps 8

Qy 96 DL$R---SLGAVF$RVTNNVL$EAPFLVGIE$SLK$STYNL-----LFCG$CGIPVGF 146
366 DIGRKKVKIG$RVFVR$NDVI--$EIMGVTE$ETGE$TNEIEAPTICPYCG$SIVKEGV 422
Db 147 HLY$TH$AAL$ARGH$CL$SD$K$KVCYLLK$TK$ATVNA$EMDIQV$P-L$EKI$E-L$EKIV 204
423 HL-----FC--$ENTL$CKPQ$MVK$SIVH$F$A$R$E$AMNIEG$F$E$K$T$ABOL$F$E-- 465
Qy 205 L$TH$N$RLK$SLM$KIL$E 219
Db 466 L$N$K$S$D$L$Y$R$I$T$K$E 480

Search completed: March 25, 2006, 08:18:21
Job time : 166 secs

```

GenCore version 5.1.7

Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 25, 2006, 08:15:45 ; Search time 24 Seconds
(without alignments)
281.425 Million cell updates/sec

Title: US-09-942-052A-728

Perfect score: 1198

Sequence: 1 MAAPLHRSCATPPRGDF.....LKSLMKILSEVTPDQSKPEN 229

Scoring table: BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 174695 seqs, 29494374 residues

Total number of hits satisfying chosen parameters: 174695

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA New:

- 1: /SIDSS/ptodata/1/pubpaa/US08 NEW PUB.pap.*
- 2: /SIDSS/ptodata/1/pubpaa/US06 NEW PUB.pap.*
- 3: /SIDSS/ptodata/1/pubpaa/US07 NEW PUB.pap.*
- 4: /SIDSS/ptodata/1/pubpaa/PCT_NEW PUB.pap.*
- 5: /SIDSS/ptodata/1/pubpaa/US09 NEW PUB.pap.*
- 6: /SIDSS/ptodata/1/pubpaa/US10 NEW PUB.pap.*
- 7: /SIDSS/ptodata/1/pubpaa/US11 NEW PUB.pap.*
- 8: /SIDSS/ptodata/1/pubpaa/US60_NEW PUB.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	86	7.2	584	7	US-11-096-568A-12723
2	86	7.2	619	7	US-11-096-568A-12723
3	86	7.2	644	7	US-11-096-568A-12721
4	80.5	6.7	1711	7	US-11-143-984A-38
5	79.5	6.6	227	7	US-11-096-568A-19014
6	79.5	6.6	278	7	US-11-096-568A-19013
7	79.5	6.6	282	7	US-11-096-568A-19012
8	79.5	6.6	306	7	US-11-172-740-327
9	79	6.6	482	7	US-11-229-371-87
10	79	6.6	482	7	US-11-229-371-177
11	79	6.6	482	7	US-11-228-923-87
12	79	6.6	482	7	US-11-228-923-177
13	79	6.6	482	7	US-11-228-875-87
14	79	6.6	482	7	US-11-228-875-177
15	79	6.6	1705	7	US-11-143-984A-37
16	77.5	6.5	462	7	US-11-087-099-12028
17	76.5	6.4	400	7	US-11-229-371-2
18	76.5	6.4	400	7	US-11-228-923-2
19	76.5	6.4	400	7	US-11-228-875-2
20	76	6.3	1001	7	US-11-072-512-2283
21	75.5	6.3	8695	7	US-11-205-109-15
22	75	6.3	295	7	US-11-172-740-323
23	75	6.3	458	7	US-11-069-642-3
24	74.5	6.2	267	7	US-11-087-099-9317
25	74.5	6.2	467	7	US-11-096-568A-11427

ALIGNMENTS

RESULT 1

US-11-096-568A-12723
; Sequence 12723, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: Therby
; CURRENT APPLICATION NUMBER: US/11/096.568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 12723
; LENGTH: 584
; TYPE: PRT
; ORGANISM: Triticum aestivum
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(584)
; OTHER INFORMATION: Ceres Seq. ID no. 14303176
US-11-096-568A-12723

Query Match 7.2%; Score 86; DB 7; Length 584;
Best Local Similarity 21.5%; Pred. No. 1.3;
Matches 47; Conservative 23; Mismatches 69; Indels 80; Gaps 9;
Qy 72 PERCAVFCQAQCHAV-----LADSVHLA-WDLRSRL-----GAVVFS 107
Db 87 PIPAEIRNCTELRAIVLRANYLOGGIPPGIGDLTHLTLIDLSNLRGAIPATIGSLTHL 146
Qy 108 RVTNVVLEAPFL-----VGIEGSLKSTY--NLPCG-----SCGIPGVFHYSTHA- 153
Db 147 RFLN---LSTNFFSGEIPNVGLTFKSSVYGNLELCGLPIQKGCRTGLGPAVLPHSD 203
Qy 154 -----ALAAARGHFCILSSDKMVCYLLKTKAIV 180
Db 204 PLSSSGVSPITNNKNTSHFLNGVVGSMSTMAIALVAVLGFL-----WVCLLSRKNGV 257
Qy 181 NASEMDIQNVLPSEKIAELKEKIVLTHNRLKSLMKILSE 219
Db 258 NYEKMDKQIVPDGAKLVTVQWNLPSYSGEIRRLLELDE 296

RESULT 2

US-11-096-568A-12722
; Sequence 12722, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:

APPLICANT: Alexandrov, Nikolai et al.
TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
FILE REFERENCE: 2750-1592PUS2
CURRENT APPLICATION NUMBER: US/11/096,568A
CURRENT FILING DATE: 2005-04-01
NUMBER OF SEQ ID NOS: 34471
SEQ ID NO 12722
LENGTH: 619
TYPE: PRT
ORGANISM: Triticum aestivum
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)..[619]
OTHER INFORMATION: Ceres Seq. ID no. 14303175
US-11-096-568A-12722

Query Match 7.2%; Score 86; DB 7; Length 619;
Best Local Similarity 21.5%; Pred. No. 1.4;
Matches 47; Conservative 23; Mismatches 69; Indels 80; Gaps 9;

QY 72 PERCAVFQCAQCHAV-----LADSVHLA-WDLRSRL-----GAVVFS 107
DB 122 PIPAEIRNCTELRAIVLRANVYQGGIPGGIGDLTHLTILDSSNLLRGAIPATIGSLTHL 181
QY 108 RVTNNVLEAPFL-----VGIEGSLKSTY--NLLFCG-----SCGIPVGFHLYSTHA- 153
DB 182 RPLN---LSTNFFSGEIPNVGVLGTFKSSSYVGNLELCGLPIQKGCRTGLGFPVAVLPHSD 238
QY 154 -----ALAAIRGHFCLSSDKMVCYLLKTKAIV 180
DB 239 PLSSSGVSPITNNKNTSHFLNGVVGISMSTMAIALVAVLGLF-----WVCLLSRKNGV 292

QY 181 NASEMDIQNVPLSEKIAELKEKIVLTHNRKLSMKILSE 219
DB 293 NYEKMDKQTVPDGAKLVTYQWNLPSYSGEIRRLLELDE 331

RESULT 3
US-11-096-568A-12721
Sequence 12721, Application US/11096568A
Publication No. US20060048240A1
GENERAL INFORMATION:
APPLICANT: Alexandrov, Nikolai et al.
TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
FILE REFERENCE: 2750-1592PUS2
CURRENT APPLICATION NUMBER: US/11/096,568A
CURRENT FILING DATE: 2005-04-01
NUMBER OF SEQ ID NOS: 34471
SEQ ID NO 12721
LENGTH: 644
TYPE: PRT
ORGANISM: Triticum aestivum
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)..[644]
OTHER INFORMATION: Ceres Seq. ID no. 14303174
US-11-096-568A-12721

Query Match 7.2%; Score 86; DB 7; Length 644;
Best Local Similarity 21.5%; Pred. No. 1.5;
Matches 47; Conservative 23; Mismatches 69; Indels 80; Gaps 9;

QY 72 PERCAVFQCAQCHAV-----LADSVHLA-WDLRSRL-----GAVVFS 107
DB 147 PIPAEIRNCTELRAIVLRANVYQGGIPGGIGDLTHLTILDSSNLLRGAIPATIGSLTHL 206
QY 108 RVTNNVLEAPFL-----VGIEGSLKSTY--NLLFCG-----SCGIPVGFHLYSTHA- 153
DB 207 RPLN---LSTNFFSGEIPNVGVLGTFKSSSYVGNLELCGLPIQKGCRTGLGFPVAVLPHSD 263
QY 154 -----ALAAIRGHFCLSSDKMVCYLLKTKAIV 180

DB 264 PLSSSGVSPITNNKNTSHFLNGVVGISMSTMAIALVAVLGLF-----WVCLLSRKNGV 317
QY 181 NASEMDIQNVPLSEKIAELKEKIVLTHNRKLSMKILSE 219
DB 318 NYEKMDKQTVPDGAKLVTYQWNLPSYSGEIRRLLELDE 356

RESULT 4
US-11-143-984A-38
Sequence 38, Application US/11143984A
Publication No. US20060014180A1
GENERAL INFORMATION:
APPLICANT: Bristol-Myers Squibb Company
TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL HUMAN PHOSPHATASES
FILE REFERENCE: D0072 DIV1
CURRENT APPLICATION NUMBER: US/11/143,984A
CURRENT FILING DATE: 2005-06-02
PRIOR APPLICATION NUMBER: US 60/256,868
PRIOR FILING DATE: 2000-12-20
PRIOR APPLICATION NUMBER: US 60/280,186
PRIOR FILING DATE: 2001-03-30
PRIOR APPLICATION NUMBER: US 60/287,735
PRIOR FILING DATE: 2001-05-01
PRIOR APPLICATION NUMBER: US 60/295,848
PRIOR FILING DATE: 2001-06-05
PRIOR APPLICATION NUMBER: US 60/300,465
PRIOR FILING DATE: 2001-06-25
NUMBER OF SEQ ID NOS: 208
SOFTWARE: PatentIn version 3.2
SEQ ID NO 38
LENGTH: 1711
TYPE: PRT
ORGANISM: Rattus norvegicus
US-11-143-984A-38

Query Match 6.7%; Score 80.5; DB 7; Length 1711;
Best Local Similarity 29.6%; Pred. No. 20;
Matches 21; Conservative 8; Mismatches 17; Indels 25; Gaps 1;

QY 36 SMEWDTQVYKGSPLGPAAGLGAEPAPQPLPSWLPQPERCAVFCQCAQCHAVLADSVHLAW 95
DB 753 SMOAGSAVNLAWPSGFLGQGA-----CHAQLSDAGHLWS 787

QY 96 DLSRSLGAVVF 106
DB 788 EQPLKLGQELF 798

RESULT 5
US-11-096-568A-19014
Sequence 19014, Application US/11096568A
Publication No. US20060048240A1
GENERAL INFORMATION:
APPLICANT: Alexandrov, Nikolai et al.
TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
FILE REFERENCE: 2750-1592PUS2
CURRENT APPLICATION NUMBER: US/11/096,568A
CURRENT FILING DATE: 2005-04-01
NUMBER OF SEQ ID NOS: 34471
SEQ ID NO 19014
LENGTH: 227
TYPE: PRT
ORGANISM: Zea mays subsp. mays
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)..(227)
OTHER INFORMATION: Ceres Seq. ID no. 12368187
NAME/KEY: misc feature
LOCATION: (122)..(122)
OTHER INFORMATION: Xaa is any aa, unknown or other

```
;
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (126)..(126)
; OTHER INFORMATION: Xaa is any aa, unknown or other
;
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (133)..(133)
; OTHER INFORMATION: Xaa is any aa, unknown or other
US-11-096-568A-19014

Query Match          6.6%; Score 79.5; DB 7; Length 227;
Best Local Similarity 24.5%; Pred. No. 1.6;
Matches 24; Conservative 11; Mismatches 40; Indels 23; Gaps 4;

Qy 10 SRCATPRGDFCGGTERAIDQASFTTSMEDTQVVGSSPLGPAGLGAEEPAAGPQLPSW 69
Db 118 SGCCXPP--TYCNFT-----YTXGTEW-TRTAAGSAP-----AGRDCSAW 154

Qy 70 LQPERCAVFOCAQCHAVLADSVHLAWDLRSLSGAVVFS 107
Db 155 GNDEDDL CYGCSCKAGVVDALKRDWKRAAIVNVVILS 192

RESULT 6
US-11-096-568A-19013
; Sequence 19013, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nickolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 19013
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(278)
; OTHER INFORMATION: Xaa is any aa, unknown or other
;
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (173)..(173)
; OTHER INFORMATION: Xaa is any aa, unknown or other
;
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (184)..(184)
; OTHER INFORMATION: Xaa is any aa, unknown or other
US-11-096-568A-19013

Query Match          6.6%; Score 79.5; DB 7; Length 278;
Best Local Similarity 24.5%; Pred. No. 2.1;
Matches 24; Conservative 11; Mismatches 40; Indels 23; Gaps 4;

Qy 10 SRCATPRGDFCGGTERAIDQASFTTSMEDTQVVGSSPLGPAGLGAEEPAAGPQLPSW 69
Db 169 SGCCXPP--TYCNFT-----YTXGTEW-TRTAAGSAP-----AGRDCSAW 205

Qy 70 LQPERCAVFOCAQCHAVLADSVHLAWDLRSLSGAVVFS 107
Db 206 GNDEDDL CYGCSCKAGVVDALKRDWKRAAIVNVVILS 243

RESULT 7
US-11-096-568A-19012
; Sequence 19012, Application US/11096568A
```

```
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nickolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 19012
; TYPE: PRT
; LENGTH: 282
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(282)
; OTHER INFORMATION: Ceres Seq. ID no. 12368185
;
; NAME/KEY: misc feature
; LOCATION: (177)..(177)
; OTHER INFORMATION: Xaa is any aa, unknown or other
;
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (181)..(181)
; OTHER INFORMATION: Xaa is any aa, unknown or other
;
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (188)..(188)
; OTHER INFORMATION: Xaa is any aa, unknown or other
US-11-096-568A-19012

Query Match          6.6%; Score 79.5; DB 7; Length 282;
Best Local Similarity 24.5%; Pred. No. 2.2;
Matches 24; Conservative 11; Mismatches 40; Indels 23; Gaps 4;

Qy 10 SRCATPRGDFCGGTERAIDQASFTTSMEDTQVVGSSPLGPAGLGAEEPAAGPQLPSW 69
Db 173 SGCCXPP--TYCNFT-----YTXGTEW-TRTAAGSAP-----AGRDCSAW 209

Qy 70 LQPERCAVFOCAQCHAVLADSVHLAWDLRSLSGAVVFS 107
Db 210 GNDEDDL CYGCSCKAGVVDALKRDWKRAAIVNVVILS 247

RESULT 8
US-11-172-740-327
; Sequence 327, Application US/11172740
; Publication No. US2006057724A1
; GENERAL INFORMATION:
; APPLICANT: MASCIA, Peter
; APPLICANT: ALEXANDROV, Nickolai
; APPLICANT: BROVER, Vyacheslav
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES AND POLYPEPTIDES ENCODED THEREBY USEFUL FOR
; TITLE OF INVENTION: PLANT CHARACTERISTICS AND PHENOTYPES
; FILE REFERENCE: 2750-1602PUS2
; CURRENT APPLICATION NUMBER: US/11/172,740
; CURRENT FILING DATE: 2005-06-30
; PRIOR APPLICATION NUMBER: 60/583,621
; PRIOR FILING DATE: 2004-06-30
; PRIOR APPLICATION NUMBER: 60/584,829
; PRIOR FILING DATE: 2004-06-30
; PRIOR APPLICATION NUMBER: 60/584,800
; PRIOR FILING DATE: 2004-06-30
; NUMBER OF SEQ ID NOS: 2523
; SEQ ID NO 327
; TYPE: PRT
; LENGTH: 306
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(306)
; OTHER INFORMATION: Ceres CLONE ID no. 1459706
;
; FEATURE:
```


RESULT 11
US-11-228-923-87
; Sequence 87, Application US/11228923
; Publication No. US20060031966A1
; GENERAL INFORMATION:
; APPLICANT: McCourt, Peter
; APPLICANT: Ghassemlian, Majid
; APPLICANT: Cutler, Sean
; APPLICANT: Bonetta, Dario
; TITLE OF INVENTION: Stress Tolerance and Delayed Senescence in Plants
; FILE REFERENCE: 22542-007C1P2CON2
; CURRENT APPLICATION NUMBER: US/11/228,923
; CURRENT FILING DATE: 2005-09-16
; PRIOR APPLICATION NUMBER: 10/229,541
; PRIOR FILING DATE: 2002-08-27
; PRIOR APPLICATION NUMBER: 10/160,764
; PRIOR FILING DATE: 2002-05-31
; PRIOR APPLICATION NUMBER: 60/294,766
; PRIOR FILING DATE: 2001-05-31
; PRIOR APPLICATION NUMBER: 60/348,909
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: 10/210,760
; PRIOR FILING DATE: 2002-08-01
; PRIOR APPLICATION NUMBER: 60/309,396
; PRIOR FILING DATE: 2001-08-01
; PRIOR APPLICATION NUMBER: 60/337,084
; PRIOR FILING DATE: 2001-12-04
; PRIOR APPLICATION NUMBER: 09/191,687
; PRIOR FILING DATE: 1998-11-13
; NUMBER OF SEQ ID NOS: 178
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 87
; LENGTH: 482
; TYPE: PRT
; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: Wigwam gene
US-11-228-923-87

Query Match	6.6%	Score 79	DB 7	Length 482
Best Local Similarity	26.3%	Pred. No. 5.1		
Matches	47	Conservative 25	Mismatches 49	Indels 58
Gaps	13			

Qy	28	IDQASFTTSEWDIQ----	VVKGSSPLCP--AGLGAEEPAAGPQLPSWLQPERCAVFOCAQ	82
Db	72	VSTQKNMEIQDKQDLYLWKGLRQLGQPFSSLDANRP-----WL-		111
Qy	83	CHAVLADSVHLAMDLSRLGAAVVFSTRVTNNVVLAEAPLVGIEGSLKSTYNLLFCGSCG-	141	
Db	112	CYWIL-HSIAL-----LGETVDDELESNAI---DFLGRCSGEGG-----YGGPGQ	154	
Qy	142	IPGVGHLYSTHA---ALAAURGHFCLSS-----DKMVCVLLATK-----ATVNASEMDIQ	188	
Db	155	LP---HLATTYAAVNAVLTTGGKALSSINREKMSCSFLREMKDTSGGFRMHMDGEMDVR	210	

RESULT 12
US-11-228-923-177
; Sequence 177, Application US/11228923
; Publication No. US20060031966A1
; GENERAL INFORMATION:
; APPLICANT: McCourt, Peter
; APPLICANT: Ghassemlian, Majid
; APPLICANT: Cutler, Sean
; APPLICANT: Bonetta, Dario
; TITLE OF INVENTION: Stress Tolerance and Delayed Senescence in Plants
; FILE REFERENCE: 22542-007CIP2CON2
; CURRENT APPLICATION NUMBER: US/11/228,923
; CURRENT FILING DATE: 2003-09-16
; PRIOR APPLICATION NUMBER: 10/229,541
; PRIOR FILING DATE: 2002-08-27

```

; PRIOR APPLICATION NUMBER: 10/160,764
; PRIOR FILING DATE: 2002-05-31
; PRIOR APPLICATION NUMBER: 60/294,766
; PRIOR FILING DATE: 2001-05-31
; PRIOR APPLICATION NUMBER: 60/348,909
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: 10/210,760
; PRIOR FILING DATE: 2002-08-01
; PRIOR APPLICATION NUMBER: 60/309,396
; PRIOR FILING DATE: 2001-08-01
; PRIOR APPLICATION NUMBER: 60/337,084
; PRIOR FILING DATE: 2001-12-04
; PRIOR APPLICATION NUMBER: 09/191,687
; PRIOR FILING DATE: 1998-11-13
; NUMBER OF SEQ ID NOS: 178
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 177
; LENGTH: 482
; TYPE: PRT
; ORGANISM: Arabidopsis sp.
US-11-228-923-177

Query Match          6.6%; Score 79; DB 7; Length 482;
Best Local Similarity 26.3%; Pred No. 5.1;
Matches 47; Conservative 25; Mismatches 49; Indels 58; Gaps 13;

Qy 28 IDQASFTTSMEDTQ---VVGSSPLGP--AGIGAEPAAGPOLPSWLQPERCAVFQCAQ 82
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 72 VSTQKYMMETIQDKLDYLMKGLQLGPOFSSLDANRP-----WL----- 111

Qy 83 CHAVLADSVHLAWLDSRSJGAVFVSRTNNVNVLEAPLVGIEGLKSTYNLLFCGSCG- 141
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 112 CYWIL-HSTAL-----JGETVDDESNAI---DFLGRCSQSEGG-----YGGPGQ 154

Qy 142 IPVGPHLYSTHA---ALAAALRGHFCLSS---DKMVCYLLKTK-----AIVNASEMDIQ 188
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 155 LP---HLATYAANVAALVTLGDKKALSSINREKMSCFLRMKDTSGGFRNHDGEMDVR 210
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

RESULT 13
US-11-228-875-87
; Sequence 87, Application US/11228875
; Publication No. US20060037108A1
; GENERAL INFORMATION:
; APPLICANT: McCourt, Peter
; APPLICANT: Ghassemlian, Majid
; APPLICANT: Cutler, Sean
; APPLICANT: Bonetta, Dario
; TITLE OF INVENTION: Stress Tolerance and Delayed Senescence in Plants
; FILE REFERENCE: 22542-007CIP2CON3
; CURRENT APPLICATION NUMBER: US/11/228,875
; CURRENT FILING DATE: 2005-09-16
; PRIOR APPLICATION NUMBER: 10/229,541
; PRIOR FILING DATE: 2002-08-27
; PRIOR APPLICATION NUMBER: 10/160,764
; PRIOR FILING DATE: 2002-05-31
; PRIOR APPLICATION NUMBER: 60/294,766
; PRIOR FILING DATE: 2001-05-31
; PRIOR APPLICATION NUMBER: 60/348,909
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: 10/210,760
; PRIOR FILING DATE: 2002-08-01
; PRIOR APPLICATION NUMBER: 60/309,396
; PRIOR FILING DATE: 2001-08-01
; PRIOR APPLICATION NUMBER: 60/337,084
; PRIOR FILING DATE: 2001-12-04
; PRIOR APPLICATION NUMBER: 09/191,687
; PRIOR FILING DATE: 1998-11-13
; NUMBER OF SEQ ID NOS: 178
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 87
; LENGTH: 482
; TYPE: PRT

```
; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: Wigwum gene
US-11-228-875-87

Query Match      6.6%; Score 79; DB 7; Length 482;
Best Local Similarity 26.3%; Pred. No. 5.1;
Matches 47; Conservative 25; Mismatches 49; Indels 58; Gaps 13;

Qy 28 IDQASFTTSMEDTQ---VVKGSSPLGP--AGLGAEEPAAGPQLPSWLQPERCAVFOCAQ 82
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
72 VSTQKYMMEIQDKQLDYLKMLKGLRQLGPFSSLDANRP-----WL-----111

Qy 83 CHAVLADSVHLAWLDSRLSGAVVFSRVTVNVVLEAPFLVGLGSLKSTYNLLFCGSCG- 141
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
112 CYWIL-HSIAL-----LGTVDDELESNAI---DFLGRQCGSEGG-----YGGPGQ 154

Qy 142 IPVGFHLYSTHA---ALAAALRGHFCCLSS---DKMVCYLLKTK-----AIVNASEMDIQ 188
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
155 LP---HLATTYAAVNALVTGGDKALSSINERKMSCLRRMKDTSGGFRMHDMGEMDVR 210

RESULT 14
US-11-228-875-177
; Sequence 177, Application US/11228875
; Publication No. US20060037108A1
; GENERAL INFORMATION:
; APPLICANT: McCourt, Peter
; APPLICANT: Ghasseman, Majid
; APPLICANT: Cutler, Sean
; APPLICANT: Bonetta, Dario
; TITLE OF INVENTION: Stress Tolerance and Delayed Senescence in Plants
; FILE REFERENCE: 22542-00/CIPZCON3
; CURRENT APPLICATION NUMBER: US/11/228,875
; CURRENT FILING DATE: 2005-09-16
; PRIOR APPLICATION NUMBER: 10/229,541
; PRIOR FILING DATE: 2002-08-27
; PRIOR FILING DATE: 2002-05-31
; PRIOR APPLICATION NUMBER: 60/294,766
; PRIOR FILING DATE: 2001-05-31
; PRIOR APPLICATION NUMBER: 60/348,909
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: 10/210,760
; PRIOR FILING DATE: 2002-08-01
; PRIOR APPLICATION NUMBER: 60/309,396
; PRIOR FILING DATE: 2001-08-01
; PRIOR APPLICATION NUMBER: 60/337,084
; PRIOR FILING DATE: 2001-12-04
; PRIOR APPLICATION NUMBER: 09/191,687
; PRIOR FILING DATE: 1998-11-13
; NUMBER OF SEQ ID NOS: 178
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 177
; LENGTH: 482
; TYPE: PRT
; ORGANISM: Arabidopsis sp.
US-11-228-875-177

Query Match      6.6%; Score 79; DB 7; Length 482;
Best Local Similarity 26.3%; Pred. No. 5.1;
Matches 47; Conservative 25; Mismatches 49; Indels 58; Gaps 13;

Qy 28 IDQASFTTSMEDTQ---VVKGSSPLGP--AGLGAEEPAAGPQLPSWLQPERCAVFOCAQ 82
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
72 VSTQKYMMEIQDKQLDYLKMLKGLRQLGPFSSLDANRP-----WL-----111

Qy 83 CHAVLADSVHLAWLDSRLSGAVVFSRVTVNVVLEAPFLVGLGSLKSTYNLLFCGSCG- 141
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
112 CYWIL-HSIAL-----LGTVDDELESNAI---DFLGRQCGSEGG-----YGGPGQ 154

Qy 142 IPVGFHLYSTHA---ALAAALRGHFCCLSS---DKMVCYLLKTK-----AIVNASEMDIQ 188
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
155 LP---HLATTYAAVNALVTGGDKALSSINERKMSCLRRMKDTSGGFRMHDMGEMDVR 210

US-09-942-052a-728.rapbn

Db 155 LP---HLATTYAAVNALVTGGDKALSSINERKMSCLRRMKDTSGGFRMHDMGEMDVR 210

RESULT 15
US-11-143-984A-37
; Sequence 37, Application US/11143984A
; Publication No. US20060014180A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL HUMAN PHOSPHATASES
; FILE REFERENCE: D0072 DIV1
; CURRENT APPLICATION NUMBER: US/11/143,984A
; CURRENT FILING DATE: 2005-06-02
; PRIOR APPLICATION NUMBER: US 60/256,868
; PRIOR FILING DATE: 2000-12-20
; PRIOR APPLICATION NUMBER: US 60/280,186
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: US 60/287,735
; PRIOR FILING DATE: 2001-05-01
; PRIOR APPLICATION NUMBER: US 60/295,848
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/300,465
; PRIOR FILING DATE: 2001-06-25
; NUMBER OF SEQ ID NOS: 208
; SOFTWARE: Patent In version 3.2
; SEQ ID NO 37
; LENGTH: 1705
; TYPE: PRT
; ORGANISM: Mus musculus
US-11-143-984A-37

Query Match      6.6%; Score 79; DB 7; Length 1705;
Best Local Similarity 25.2%; Pred. No. 28;
Matches 26; Conservative 17; Mismatches 30; Indels 30; Gaps 3;

Qy 36 SMENDTOVVKGSSPLGPAAGLGAEEPAAGPQLPSWLQPERCAVFOCAOCHAVLADSVHLAW 95
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
753 SMQAGSAVNLAWPSPGLGQGT-----CHAQLSDAGHLWSW 787

Qy 96 DLRSRLGAVVFSRVTVNVVLEAPFLVGIE---GSLKSTYNLL 135
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
788 EQPLSLGQDL--MLRNLPHTVSLVKCRAGPLQASTHPLV 828

Search completed: March 25, 2006, 08:18:51
Job time : 25 secs
```